

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



Cited in Patent  
Previous Office  
Actions

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT) *Actions*

<p>(51) International Patent Classification 5 : C12P 19/34, C12N 5/10, 7/01 C12N 9/10, 9/12, 15/00 C12N 15/10, 15/82, 15/85 C07H 15/12</p>		A1	<p>(11) International Publication Number: <b>WO 92/13089</b></p> <p>(43) International Publication Date: 6 August 1992 (06.08.92)</p>
<p>(21) International Application Number: PCT/US92/00277</p> <p>(22) International Filing Date: 16 January 1992 (16.01.92)</p> <p>(30) Priority data: 642,330 17 January 1991 (17.01.91) US</p> <p>(60) Parent Application or Grant (63) Related by Continuation US 642,330 (CIP) Filed on 17 January 1992 (17.01.92)</p> <p>(71) Applicants (for all designated States except US): THE GENERAL HOSPITAL CORPORATION [US/US]; Fruit Street, Boston, MA 02114 (US). THE PRESIDENT AND FELLOWS OF HARVARD COLLEGE [US/US]; 17 Quincy Street, Cambridge, MA 02138 (US).</p>		<p>(72) Inventors; and (75) Inventors/Applicants (for US only) : HASELOFF, James [AU/US]; BRAND, Andrea [US/US]; 26 Maple Avenue, Cambridge, MA 02139 (US). PERRIMON, Norbert [FR/US]; 76 Regent Circle, Brookline, MA 02146 (US). GOODMAN, Howard, M. [US/US]; 10 The Ledges Road, Newton, MA 02159 (US).</p> <p>(74) Agents: GOLDSTEIN, Jorge, A. et al.; Sterne, Kessler, Goldstein &amp; Fox, 1225 Connecticut Avenue, N.W., Suite 300, Washington, DC 20036 (US).</p> <p>(81) Designated States: AT (European patent), AU, BE (European patent), CA, CH (European patent), CS, DE (European patent), DK (European patent), ES (European patent), FI, FR (European patent), GB (European patent), GR (European patent), HU, IT (European patent), JP, KR, LU (European patent), MC (European patent), NL (European patent), NO, RU, SE (European patent), US.</p> <p>Published <i>With international search report.</i></p>	
<p>(54) Title: CELL ABLATION USING TRANS-SPlicing RIBOZYMES</p> <p>(57) Abstract</p> <p>The design of new ribozymes capable of self-catalyzed trans-splicing which are based upon the catalytic core of a Group I intron are described. Using this design, it is possible to construct ribozymes capable of efficiently splicing a new 3' exon sequence into any chosen target RNA sequence in a highly precise manner. A method of cell ablation is also described that provides a toxic product to a host cell <i>in vivo</i> in a targetted, regulated manner utilizing novel trans-splicing ribozymes of the invention. Inactive pro-ribozyme forms are also described.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FI	Finland	ML	Mali
AU	Australia	FR	France	MN	Mongolia
BB	Barbados	GA	Gabon	MR	Mauritania
BE	Belgium	GB	United Kingdom	MW	Malawi
BF	Burkina Faso	GN	Guinea	NL	Netherlands
BG	Bulgaria	GR	Greece	NO	Norway
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IE	Ireland	RO	Romania
CA	Canada	IT	Italy	RU	Russian Federation
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland	KR	Republic of Korea	SN	Senegal
CI	Côte d'Ivoire	LI	Liechtenstein	SU	Soviet Union
CM	Cameroon	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
DE	Germany	MC	Monaco	US	United States of America
DK	Denmark	MG	Madagascar		
ES	Spain				

-1-

**TITLE OF THE INVENTION**

5

**CELL ABLATION USING  
TRANS-SPlicing RIBOZYMES**

**Field of the Invention**

10

The present invention is directed to novel trans-splicing ribozymes and methods of cell ablation using these ribozymes.

15

**BRIEF DESCRIPTION OF THE BACKGROUND ART**

**I. Group I Introns**

20

RNA molecules with catalytic activity are called ribozymes or RNA enzymes (Cech, T.R., *Ann. Rev. Biochem.* 59:543-568 (1990)). The *Tetrahymena thermophila* precursor rRNA contains an intron (a ribozyme) capable of catalyzing its own excision. This ribozyme is one of a class of structurally related Group I introns.

25

The splicing activity of the modified *T. thermophila* intron requires the presence of a guanosine cofactor and a divalent cation, either  $Mg^{++}$  or  $Mn^{++}$ , and occurs via two sequential transesterification reactions (Figure 1). First, a free guanosine is bound to the ribozyme and its 3' hydroxyl group is positioned to attack the phosphorus atom at the 5' splice site. The guanosine is covalently attached to the intron sequence and the 5' exon is released. Second, the phosphodiester bond located at the 3' splice site undergoes attack from the newly freed 3' hydroxyl group of the 5' exon, resulting in production

30

35

- 2 -

of the ligated exon sequences. The excised intron subsequently undergoes a series of transesterification reactions, involving its 3' hydroxyl group and internal sequences, resulting in the formation of shortened circular forms.

These successive reactions are chemically similar and appear to occur at a single active site. The reactions of self-splicing are characterized by the formation of alternative RNA structures as differing RNA chains are each brought to form similar conformations around the highly conserved intron. Splicing requires the alignment of the intron-exon junctions across a complementary sequence termed the "internal guide sequence" or IGS.

The first cleavage at the 5' splice site requires the formation of a base-paired helix (P1) between the IGS and sequences adjacent the splice site. The presence of a U:G "wobble" base-pair within this helix defines the phosphodiester bond that will be broken in the catalytic reaction of the ribozyme. After cleavage of this bond, a portion of the P1 helix is displaced and a new helix, P10, is formed due to complementarity between the IGS and sequences adjacent the 3' splice site. An invariant guanosine residue precedes the phosphodiester at the 3' splice site, similar to the portion of the P1 sequence that it is displacing. Thus, ligation of the exons occurs in a reverse of the first cleavage reaction but where new exon sequences have been substituted for those of the intron. It may be noted that intron circularization reactions subsequent to exon ligation also involve base-pairing of 5' sequences across the IGS, and attack mediated by the 3' hydroxyl group of the intron's terminal guanine residue (Been, M.D. et al., "Selection Of Circularizaton Sites In A Group I IVS RNA Requires

- 3 -

Multiple Alignments Of An Internal Template-Like Sequence," *Cell* 50:951 (1987)).

## II. Catalytic Activities

In order to better define the structural and catalytic properties of the Group I introns, exon sequences have been stripped from the "core" of the *T. thermophila* intron. Cech, T.R. et al., WO 88/04300, describes at least three catalytic activities possessed by the *Tetrahymena* intron ribozyme: (1) a dephosphorylating activity, capable of removing the 3' terminal phosphate of RNA in a sequence-specific manner, (2) an RNA polymerase activity (nucleotidyl transferase), capable of catalyzing the conversion of oligoribonucleotides to polyribonucleotides, and (3) a sequence-specific endoribonuclease activity.

Isolated ribozyme activities can interact with substrate RNAs in *trans*, and these interactions characterized. For example, when truncated forms of the intron are incubated with sequences corresponding to the 5' splice junction, the site undergoes guanosine-dependent cleavage in mimicry of the first step in splicing. The substrate and endoribonucleolytic intron RNAs base-pair to form helix P1, and cleavage occurs after a U:G base-pair at the 4th-6th position. Phylogenetic comparisons and mutational analyses indicate that the nature of the sequences immediately adjacent the conserved uracil residue at the 5' splice site are unimportant for catalysis, provided the base-pairing of helix P1 is maintained (Doudna, J.A. et al., *Proc. Natl. Acad. Sci. USA* 86: 7402-7406 (1989)).

The sequence requirements for 3' splice-site selection appear to lie mainly within the structure of the intron itself, including helix P9.0 and the following guanosine residue which delineates the 3' intron boundary. However, flanking sequences within the

- 4 -

5                   3' exon are required for the formation of helix P10 and efficient splicing, as shown by mutational analysis (Suh, E.R. et al., Mol. Cell. Biol. 10:2960-2965 (1990)). In addition, oligonucleotides have been  
10                 ligated in *trans*, using a truncated form of the intron, and "external" guide sequence and oligonucleotides which had been extended by a 5' guanosine residue. The substrate oligonucleotides corresponding to 3' exon sequences were aligned solely by the formation of P10-like helices on an external template, prior to ligation (Doudna, J.A. et al., Nature 339:519-522 (1989)).

15                 The cleavage activity of ribozymes has been targeted to specific RNAs by engineering a discrete "hybridization" region into the ribozyme, such hybridization region being capable of specifically hybridizing with the desired RNA. For example, Gerlach, W.L. et al., EP 321,201, constructed a ribozyme containing a sequence complementary to a target RNA. Increasing the length of this complementary sequence increased the affinity of this sequence for the target. However, the hybridizing and cleavage regions of this ribozyme were integral parts of each other. Upon hybridizing to the target RNA through the complementary regions, the catalytic region 20                 of the ribozyme cleaved the target. It was suggested that the ribozyme would be useful for the inactivation or cleavage of target RNA *in vivo*, such as for the treatment of human diseases characterized by the production of a foreign host's RNA. However, ribozyme-directed *trans*-splicing, (as opposed to *trans*-cleavage) 25                 was not described or suggested.

30                 The endoribonuclease activities (the cleavage activities) of various naturally-occurring ribozymes have been extensively studied. Analysis of the structure and sequence of these ribozymes has indicated 35                 that certain nucleotides around the cleavage site are

- 5 -

highly conserved but flanking sequences are not so conserved. This information has led to the design of novel endoribonuclease activities not found in nature. For example, Cech and others have constructed novel ribozymes with altered substrate sequence specificity (Cech, T.R. et al., WO 88/04300; Koizumi, M. et al., FEBS Lett. 228:228-230 (1988); Koizumi, M. et al., FEBS Lett. 239:285-288 (1988); Haseloff, J. et al., Nature 334:585-591 (1987); and Heus, H.A. et al., Nucl. Acids Res. 18:1103-1108 (1990)). From early studies of the self-cleaving plant viroids and satellite RNAs (Buzayan, J.M. et al., Proc. Natl. Acad. Sci. USA 83:8859-8862 (1986), guidelines for the design of ribozymes that are capable of cleaving other RNA molecules in *trans* in a highly sequence specific have been developed (Haseloff, J. et al., Nature 334:585-591 (1988)). However, these constructs were unable to catalyze efficient, targeted *trans*-splicing reactions.

The joining of exons contained on separate RNAs, that is, *trans*-splicing, occurs in nature for both snRNP-mediated and self-catalyzed group I and group II introns. In trypanosome and *Caenorhabditis elegans* mRNAs, common 5' leader sequences are transcribed from separate genes and spliced to the 3' portions of the mRNAs (Agabian, N., Cell 61:1157-1160 (1990); Hirsh, D. et al., Mol. Biol. Rep. 14:115 (1990)). These small "spliced leader" RNAs (sLRNAs) consist of the 5' exon fused to sequences that can functionally substitute for U1 snRNA in mammalian snRNP-splicing extracts.

Also, both the group I and group II self-splicing introns are capable of exon ligation in *trans* in artificial systems (Been, M.D. et al., Cell 47:207-216 (1986); Galloway-Salvo, J.L. et al., J. Mol. Biol. 211:537-549 (1990); Jacquier, A. et al., Science 234:1099-1194 (1986); and Jarrell, K.A. et al., Mol. Cell Biol. 8:2361-2366 (1988)). *Trans*-splicing occurs

- 6 -

in vivo for group II introns in split genes of chloroplasts (Kohchi, T. et al., *Nucl. Acids Res.* 16:10025-10036 (1988)), and has been shown for a group I intron in an artificially split gene in *Escherichia coli* (Galloway-Salvo, J.L. et al., *J. Mol. Biol.* 211:537-549 (1990)). In the latter case, a bacteriophage T4 thymidylate synthase gene (td) containing a group I intron was divided at the loop connecting the intron helix P6a. Transcripts of the td gene segments were shown to undergo trans-splicing in vitro, and to rescue dysfunctional *E. coli* host cells. Known base-pairings (P3, P6 and P6a) and possible tertiary interactions between the intron segments, allowed correct assembly and processing of the gene halves.

In vitro, the *Tetrahymena* ribozyme is capable of catalyzing the trans-splicing of single-stranded model oligoribonucleotide substrates. Four components were necessary: ribozyme, 3' single-stranded RNA, 5' exon and GTP. A shortened form of the *Tetrahymena* ribozyme (L-21 *Scal* IVS RNA), starting at the internal guide sequence and terminating at U<sub>409</sub> has been used in such a reaction (Flanagan, J.B. et al., *J. Cell. Biochem. (Supp.)* 12 part D:28 (1988)). Attack by GTP at the 5' splice site released the 5' exon which was then ligated by the ribozyme to the 3' exon in a transesterification reaction at the 3' splice site.

The in vivo use of ribozymes as an alternative to the use of antisense RNA for the targeting and destruction of specific RNAs has been proposed (Gerlach, W.L. et al., EP321,201; Cotten, M., *Trends Biotechnol.* 8:174-178 (1990); Cotten, M. et al., *EMBO J.* 8:3861-3866 (1989); Sarver, N. et al., *Science* 247:1222-1225 (1990)). For example, expression of a ribozyme with catalytic endonucleolytic activity towards an RNA expressed during HIV-1 infection has

- 7 -

been suggested as a potential therapy against human immunodeficiency virus type 1 (HIV-1) infection (Sarver, N. et al., *Science* 247:1222-1225 (1990); Cooper, M., *CDC AIDS Weekly*, April 3, 1989, page 2; Rossi, J.J., Abstract of Grant No. 1R01AI29329 in Dialog's Federal Research in Progress File 265). However, such attempts have not yet been successful.

In a study designed to investigate the potential use of ribozymes as therapeutic agents in the treatment of human immunodeficiency virus type 1 (HIV-1) infection, ribozymes of the hammerhead motif (Hutchins, C.J. et al., *Nucl. Acids Res.* 14:3627 (1986); Keesee, P. et al., in *Viroids and Viroid-Like Pathogens*, J.S. Semancik, ed., CRC Press, Boca Raton, FL, 1987, pp. 1-47) were targeted to the HIV-1 gag transcripts. Expression of the gag-targeted ribozyme in human cell cultures resulted in a decrease (but not a complete disappearance of) the level of HIV-1 gag RNA and in antigen p24 levels (Sarver, N. et al., *Science* 247:1222-1225 (1990)). Thus, the medical effectiveness of Sarver's ribozyme was limited by its low efficiency since any of the pathogen's RNA that escapes remains a problem for the host.

Another problem with *in vivo* ribozyme applications is that a high ribozyme to substrate ratio is required for ribozyme inhibitory function in nuclear extracts and it has been difficult to achieve such ratios.

Cotton et al. achieved a high ribozyme to substrate ratio by microinjection of an expression cassette containing a ribozyme-producing gene operably linked to a strong tRNA promoter (a polymerase III promoter) in frog oocytes, together with substrate RNA that contains the cleavage sequence for the ribozyme (Cotton, M. et al., *EMBO J.* 8:3861-3866 (1989)). However, microinjection is not an appropriate method of delivery in multicellular organisms.

- 8 -

5           The *in vivo* activity of ribozymes designed against mRNA coding for *Escherichia coli*  $\beta$ -galactosidase has been reported (Chuat, J.-C. et al., *Biochem. Biophys. Res. Commun.* 162:1025-1029 (1989)). However, this activity was only observed when the ribozyme and target were transfected into bacterial cells on the same molecule. Ribozyme activity was inefficient when targeted against an mRNA transcribed from a bacterial 10 F episome that possessed the target part of the  $\beta$ -galactosidase gene.

10           Thus, current technological applications of ribozyme activities are limited to those which propose to utilize a ribozyme's cleavage activity to destroy the activity of a target RNA. Unfortunately, such applications often require complete destruction of all 15 target RNA molecules, and/or relatively high ribozyme:substrate ratios to ensure effectiveness and this has been difficult to achieve. Most importantly, the modified ribozymes of the art are not capable of 20 efficient, directed trans-splicing.

25           Accordingly, a need exists for the development of highly efficient ribozymes and ribozyme expression systems. Especially, the art does not describe an effective means in which to destroy an existing RNA sequence or to alter the coding sequence of an existing RNA by the trans-splicing of a new RNA sequence into a host's RNA.

#### SUMMARY OF THE INVENTION

30           Recognizing the potential for the design of novel ribozymes, and cognizant of the need for highly efficient methods to alter the genetic characteristics of higher eukaryotes *in vivo*, the inventors have 35 investigated the use of ribozymes to alter the genetic information of native RNA's *in vivo*. These efforts

- 9 -

have culminated in the development of highly effective trans-splicing ribozymes, and guidelines for the engineering thereof.

According to the invention, there is first provided an RNA or DNA molecule, such molecule encoding a trans-splicing ribozyme, such ribozyme being capable of efficiently splicing a new 3' exon sequence into any chosen target RNA sequence in a highly precise manner, *in vitro* or *in vivo*, and such molecule being novel in the ability to accommodate, any chosen target RNA or 3' exon sequences, and in the addition of a complementary sequence which enhances the specificity of such ribozyme.

According to the invention, there is also provided an RNA or DNA molecule, such molecule encoding a ribozyme, the sequence for such ribozyme being a fusion RNA, such fusion RNA providing a first RNA sequence that is sufficient for targeting such ribozyme to hybridize to a target RNA, and further a second RNA sequence, such second RNA sequence capable of being transposed into the target RNA, and such second RNA sequence encoding an RNA sequence foreign to the targeted RNA sequence.

According to the invention, there is further provided an RNA or DNA molecule, such molecule encoding a ribozyme, the sequence for such ribozyme being a fusion RNA as described above, the first RNA sequence provided by the fusion RNA being a sequence for targeting such RNA molecule to hybridize to GAL4 RNA, and the second RNA sequence of the fusion RNA providing the coding sequence of the A chain of diphtheria toxin (DTA).

According to the invention, there is also provided an RNA or DNA molecule, such molecule encoding a conformationally disrupted ribozyme of the invention, a pro-ribozyme, such pro-ribozyme being substrate-

- 10 -

activated, that is, such pro-ribozyme possessing negligible or no self-cleavage or trans-splicing activity, until being reactivated by specific interaction with target RNA.

5 According to the invention, there is further provided an RNA or DNA molecule containing a ribozyme or pro-ribozyme expression cassette, such cassette being capable of being stably maintained in a host, or inserted into the genome of a host, and such cassette providing the sequence of a promoter capable of functioning in such host, operably linked to the sequence of a ribozyme or pro-ribozyme of the invention.

10 According to the invention, there is further provided an RNA or DNA molecule containing a ribozyme or pro-ribozyme expression cassette, such cassette being capable of being stably inserted into the genome of a host, such ribozyme expression cassette providing the sequence of a GAL4-responsive promoter operably linked to the sequence of a ribozyme or pro-ribozyme of the invention.

15 According to the invention, there is further provided a method for in-vitro trans-splicing, such method comprising the steps of (1) providing a ribozyme or pro-ribozyme of the invention and an appropriate substrate for such ribozyme in vitro, (2) further providing in vitro reaction conditions that promote the desired catalytic activity of such ribozyme or pro-ribozyme; and (3) allowing such ribozyme or pro-ribozyme to react with such substrate under such conditions.

20 According to the invention, there is further provided a method for in vivo trans-splicing, such method comprising the steps of (1) providing an RNA or DNA molecule of the invention to a host cell, (2) expressing the ribozyme or pro-ribozyme encoded by such molecule in such host cell, (3) expressing a substrate

- 11 -

of such ribozyme or pro-ribozyme in such host cell, and (4) allowing such ribozyme or pro-ribozyme to react with such substrate in such host cell.

According to the invention, there is further provided a method for inactivating the activity of a target RNA, such method comprising (1) providing a ribozyme or pro-ribozyme of the invention, such ribozyme or pro-ribozyme being catalytically active against such target RNA, (2) providing such target RNA, and (3) providing conditions that allow such ribozyme or pro-ribozyme to express its catalytic activity towards such target RNA.

According to the invention, there is further provided a method for providing a desired genetic sequence to a host cell *in vivo*, such method comprising (1) providing a ribozyme or pro-ribozyme of the invention to a desired host cell, such ribozyme or pro-ribozyme being catalytically active against a target RNA in such host cell, (2) providing such ribozyme or pro-ribozyme encoding such desired genetic sequence, and (3) providing conditions that allow such ribozyme or pro-ribozyme to trans-splice such desired genetic sequence into the sequence of the target RNA.

According to the invention, there is further provided a method for cell ablation in multicellular plants and animals, such method comprising providing a ribozyme or pro-ribozyme of the invention to a any host cell, and especially into a fertilized embryonic host cell, such ribozyme or pro-ribozyme encoding the sequence of a gene toxic to such host cell and such ribozyme or pro-ribozyme being capable of trans-splicing with a desired target in such host cell.

According to the invention, there is further provided a method for engineering male or female sterility in agronomically important plant species, such method comprising the ablation of any cell

- 12 -

necessary for fertility using a ribozyme or pro-ribozyme of the invention.

According to the invention, there is further provided a method of immunizing plants against plant pathogens, such method comprising the construction of transgenic plants capable of expressing a plant pathogen-specific fusion ribozyme or pro-ribozyme of the invention, and such ribozyme or pro-ribozyme being capable of ablating any host cell infected with such pathogen.

According to the invention, there is further provided a transformed, pathogen-resistant microorganism, such microorganism being resistant to a desired pathogen, such microorganism being transformed with a ribozyme or pro-ribozyme of the invention and such ribozyme or pro-ribozyme providing a catalytic activity that targets a nucleic acid molecule expressed by such pathogen.

According to the invention, there is further provided a viral pathogen capable of delivering a desired ribozyme or pro-ribozyme activity to a desired host, such ribozyme or pro-ribozyme activity being delivered by a ribozyme or pro-ribozyme of the invention.

#### DESCRIPTION OF THE FIGURES

Figure 1 is a diagram of the mechanism of ribozyme splicing of the group I intron.

Figure 2 is a diagram of structure of the (A) *Tetrahymena thermophila* rRNA intron; (B) Target mRNA and trans-splicing ribozyme or pro-ribozyme of the invention.

Figure 3(A) is a diagram of the design of a CAT-LacZ  $\alpha$ -peptide trans-splicing ribozyme; (B) is the complete DNA coding sequence of the CAT-LacZ ribozyme.

- 13 -

5 Figure 4 presents the sequences of cucumber mosaic virus (CMV) RNA 4 trans-splicing ribozymes. A: virus RNA target sequences; B: Oligonucleotide target sequences; C: CMV RNA4 - diphtheria toxin A-chain trans-splicing ribozymes.

10 Figure 5 is a comparison of cucumber mosaic virus 3/4 sequences.

15 Figure 6(A) is a diagram of the design of a Gal4-Diphtheria toxin A (DTA) trans-splicing ribozyme; (B) is the complete coding sequence of the Gal4-DTA ribozyme with the isoleucine substitution.

20 Figure 7 is a diagram of the P-element mediated "enhancer-trapping" method for expression of Gal4 protein.

25 Figure 8 presents a partial sequence of wild-type DTA and DTA 3' exon mutants.

30 Figure 9 is a map of pGaTB and pGaTN.

35 Figure 10 is a map of pUAST.

Figure 11 is a cuticle preparation of a *Drosophila* embryos expressing a Gal4-DTA trans-splicing ribozyme.

Figure 12 presents the rationale for "proto-ribozyme" design. Arrows show sites of ribozyme cleavage, "antisense" regions are shown in black, catalytic domains are shown with radial shading, and 3' "exon" sequences are shown with light shading. In the absence of the target mRNA, trans-splicing ribozymes may transiently base-pair, and react with heterologous sequences (including their own). In addition, scission at the "3' exon" junction will occur. Inactive "proto-ribozymes" are constructed to contain extra self-complementary sequences which cause the catalytic center of the ribozyme to be mis-folded. Active ribozymes are only formed after base-pairing with the intended target mRNA - and consequent displacement of the intervening secondary structure.

- 14 -

Figure 13 shows the sequence and predicted secondary structure of the CAT-LacZ trans-splicing ribozyme. Ribozyme "core" sequences are shaded (after Cech, Gene 73:259-271 (1988)). Helices P8 are shown for the unmodified ribozyme and pro-ribozymes 1 and 2, with 13 and 18 nucleotides, respectively, of sequence complementary to the "antisense" region (highlighted).

Figure 14 shows (1) active CAT-LacZ ribozyme shown schematically, with "antisense", ribozyme domain with helix P8 and 3' "exon" sequences; (2) (a) inactive CAT-LacZ pro-ribozyme 2 shown with base-pairing between sequences in the modified helix P8 and the "anti-sense" region; and (b) the active pro-ribozyme, after base-pairing with the CAT mRNA, displacement of the helix P8 - "antisense" pairing, and re-formation of helix P8.

Figure 15 shows stability of CAT-LacZ pro-ribozyme transcripts. Plasmids containing the CAT-LacZ ribozyme and pro-ribozyme sequences were cleaved with EcoRI and transcribed using T7 or SP6 RNA polymerase and [<sup>32</sup>P]UTP. Radiolabeled transcripts were fractionated by 5% polyacrylamide gel electrophoresis in 7M urea and 25% formamide, and autoradiographed. The ribozyme transcripts underwent extensive hydrolysis, primarily at the "3' exon" junction. The pro-ribozyme forms were markedly less reactive.

Figure 16 shows endoribonuclease activity of CAT-LacZ pro-ribozymes. Plasmids containing CAT-LacZ ribozyme and pro-ribozyme sequences were cleaved, with ScaI, and transcribed with T7 or SP6 RNA polymerase. Transcripts were incubated for 30' at 37°C, 45°C and 50°C in 40 mM Tris-HCl pH 7.5, 6 mM MgCl<sub>2</sub>, 2 mM spermidine, 10 mM NaCl, 2 mM GTP with radiolabeled CAT RNA, transcribed using T7 RNA polymerase from plasmid cut with PvuII. Products were fractionated by 5% polyacrylamide gel electrophoresis in 7M urea and 25% formamide, and autoradiographed. RNA mediated cleavage

- 15 -

if the 173 nt (nucleotides) CAT RNA produces 5' and 3' fragments of 76 nt and 97 nt, respectively.

Figure 17 shows the "wild-type" and modified helices P8 used for pro-ribozyme design with possible base-pairs indicated in schematic form. Those bases which are complementary to the "anti-sense" portion of the corresponding pro-ribozyme, are shown in bold type. The number of complementary bases is listed next to each helix. The helices are ordered by the stability of the corresponding pro-ribozyme transcripts, as measured by the degree of "3' exon" hydrolysis during in vitro transcription.

Figure 18 shows the stability of GAL4-DTA pro-ribozymes. Plasmids containing ribozyme and pro-ribozyme sequences were linearized with *Xba*I and transcribed using T7 RNA polymerase. Transcripts were incubated for 60' at 50°C in 40 mM Tris-HCl pH 7.5, 6 mM MgCl<sub>2</sub>, 2 mM spermidine, 10 mM NaCl, 1 mM GTP, were fractionated by 5% polyacrylamide gel electrophoresis in 7M urea and 25% formamide, and autoradiographed. Ribozyme transcripts are extensively hydrolysed under these conditions, while pro-ribozyme 1 is less so and pro-ribozyme 2 is stable.

#### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

##### I. Definitions

In the description that follows, a number of terms used in recombinant DNA (rDNA) technology are extensively utilized. In order to provide a clear and consistent understanding of the specification and claims, including the scope to be given such terms, the following definitions are provided.

Ribozyme. An RNA molecule that inherently possesses catalytic activity.

- 16 -

Trans-splice. A form of genetic manipulation whereby a nucleic acid sequence of a first polynucleotide is co-linearly linked to or inserted into the sequence of a second polynucleotide, in a manner that retains the 3'→ 5' phosphodiester linkage between such polynucleotides. By "directed" trans-splicing or "substrate-specific" trans-splicing is meant a trans-splicing reaction that requires a specific specie of RNA as a substrate for the trans-splicing reaction (that is, a specific specie of RNA in which to splice the transposed sequence). Directed trans-splicing may target more than one RNA specie if the ribozyme or pro-ribozyme is designed to be directed against a target sequence present in a related set of RNAs.

Target RNA. An RNA molecule that is a substrate for the catalytic activity of a ribozyme or pro-ribozyme of the invention.

Expression Cassette. A genetic sequence that provides sequences necessary for the expression of a ribozyme or pro-ribozyme of the invention.

Stably. By "stably" inserting a sequence into a genome is intended insertion in a manner that results in inheritance of such sequence in copies of such genome.

Operable linkage. An "operable linkage" is a linkage in which a sequence is connected to another sequence (or sequences) in such a way as to be capable of altering the functioning of the sequence (or sequences). For example, by operably linking a ribozyme or pro-ribozyme encoding sequence to a promoter, expression of the ribozyme or pro-ribozyme encoding sequence is placed under the influence or control of that promoter. Two nucleic acid sequences (such as a ribozyme or pro-ribozyme encoding sequence and a promoter region sequence at the 5' end of th

- 17 -

5 encoding sequence) are said to be operably linked if induction of promoter function results in the transcription of the ribozyme or pro-ribozyme encoding sequence and if the nature of the linkage between the two sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the expression regulatory sequences to direct the expression of the ribozyme. Thus, a promoter region would be operably linked to a nucleic acid sequence if the promoter were capable of effecting 10 the synthesis of that nucleic acid sequence.

## II. Engineering of the Ribozyme of the Invention

15 The trans-splicing ribozymes, pro-ribozymes and methods of the invention provide, for the first time, a ribozyme capable of directed trans-splicing into any RNA sequence, and especially into mature (non-intron-containing) mRNA. The trans-splicing ribozyme as described herein, with its extended complementarity to the target, greatly differs from *T. thermophila* derived endoribonuclease activities described in the art. The 20 additional complementarity of the ribozymes of the invention confers increased affinity and specificity for the target and the complementarity is not an integral part of the catalytic activity. In addition, cleavage occurs efficiently and precisely in the 25 absence of denaturants and at high concentrations of Mg<sup>++</sup>.

30 The guidelines described herein for the design of trans-splicing ribozymes are conservative, based on the well characterized properties of group I self-splicing introns and are meant to provide a general scheme for the design of any directed trans-splicing ribozyme. Accordingly, the guidelines presented herein are not 35 limited to the group I intron of the *T. thermophila*

- 18 -

pre-mRNA and may be used by one of skill in the art to design a ribozyme of the invention with the group I introns using such guidelines and knowledge in the art.

5 The native *T. thermophila* ribozyme (the intron sequence) is located from base 53 to base 465 in the sequence below of the *T. thermophila* extrachromosomal rDNA:

TGACGCAATT CAACCAAGCG CGGGTAAACG GCAGGAGTAA CTATGACTCT  
10 CTAAATAGCA ATATTTACCT TTGGAGGGAA AAGTTATCAG GCATGCACCT  
CCTAGCTAGT CTTTAAACCA ATAGATTGCA TCGGTTAAA AGGCAAGACC  
15 GTCAAATTGC GGGAAAGGGG TCAACAGCCG TTCAGTACCA AGTCTCAGGG  
GAAACTTTGA CATGGCCTTG CAAAGGGTAT GGTAATAAGC TGACGGACAT  
20 GGTCTTAACC ACGCAGCCAA GTCCTAAGTC AACAGATCTT CTGTTGATAT  
GGATGCAGTT CACAGACTAA ATGTCGGTCG GGGAAAGATGT ATTCTTCTCA  
TAAGATATAG TCGGACCTCT CCTTAATGGG AGGTAGCGGA TGAATGGATG  
25 CAACACTGGA GCCGCTGGGA ACTAATTGT ATGCGAAAGT ATATTGATTA  
GTGGAGT ACTCGTAAGG TAGCCAAATG CCTCGTCATC TAATTAGTGA  
CGCGCATGAA TGGATTA [SEQ ID NO.1]

30 (Kan, N.C. et al., *Nucl. Acids Res.* 10:2809-2822 (1982)).

35 As described herein, the directed trans-splicing ribozymes of the invention are engineered using the catalytic core of this intron. The intron, and its catalytic core can be isolated by methods known in the art. The catalytic core of the intron, that is, the truncated intron, differs from the full-length intron only in that it is truncated at the *Sca*I site, thus removing the last five nucleotides of the intron. The truncated intron RNA may be prepared by techniques known in the art or may be purchased commercially in 40 kit form from commercial sources such as, for example, product #72000 from US Biochemical, Cleveland, OH

- 19 -

5 (RNAzyme™ Tet 1.0 Kit). This US Biochemical kit provides ribozym and the protocol for th use of the ribozyme. Transcribed Tet.1 cDNA may be used as the substrate for polymerase chain reaction (PCR) mutagenesis as described below, to produce a synthetic trans-splicing enzyme.

10 Substrate specificity of the ribozyme of the invention, that is, the ability of the ribozyme to "target" a specific RNA as a substrate, is conferred by fusing complementary sequences specific to the target (substrate) RNA to the 5' terminus of the ribozyme.

15 Directed trans-splicing specificity of the ribozyme of the invention, that is, specificity in trans-splicing a desired foreign sequence of interest with the sequence of a target RNA, is conferred by providing a new 3' exon at the 3' terminus of the ribozyme. Details of the design are further provided below.

20 To alter the structural and catalytic properties of the Group I introns, exon sequences replace the flanking sequence of such introns so that only the catalytic core of the intron, the ribozyme, remains. The resulting modified ribozyme can interact with substrate RNAs in trans. When truncated forms of the 25 intron (i.e., the catalytic "core," i.e. truncated at the *Scal* site, removing the last five nucleotides of the intron) are incubated with sequences corresponding to the 5' splice junction of the native ribozyme, the site undergoes guanosine-dependent cleavage in mimicry 30 of the first step in splicing.

Engineering of the ribozymes of the invention requires consideration of the four guidelines that follow.

35 First, a splice site must be chosen within the targ t RNA. In the final trans-splicing complex, only the 5' portion of the P1 duplex is contributed by the

- 20 -

target RNA. Only a single conserv d residu , uracil, is required immediately 5' of the intended splice site. This is the sole sequence requirement in the target RNA. There is no inate structure required of the target RNA. Mature mRNA may be targeted and the trans-splicing reaction performed in the cell's cytoplasm rather than in the nucleus against pre-mRNA. This obviates the need for high concentrations of ribozyme in a cell's nucleus.

Second, having chosen a particular target sequence, compensating sequence changes must be added to the 5' section of the ribozyme in order to allow the formation of a suitable helix P1 between the target and ribozyme RNAs. It is highly desired is that the helix P1 should contain a U:G base-pair at the intended 5' splice site, and should be positioned at the 4th, 5th (preferred) or 6th position from the base of the helix (Doudna, J.A., et al., "RNA Structure, Not Sequence Determines The 5' Splice-Site Specificity of a Group I Intron," Proc. Natl. Acad. Sci. USA 86:7402-7406 (1989), incorporated herein by reference). For the native *T. thermophila* intron, P1 extends for an additional 3 base pairs past the intended 5' splice site, and, in a preferred embodiment, this is maintained in the trans-splicing ribozyme of the invention. For trans-splicing to be efficient, the substrate and endoribonucleolytic intron RNAs must base-pair to form helix P1, with a resulting wobble U:G base-pair. Cleavage of the target RNA occurs at the phosphodiester bond immediately 3' to (after the) U:G base-pair. Phylogenetic comparisons and mutational analyses indicate that the nature of the sequences immediately adjacent the conserved uracil residue at the 5' splic site are unimportant f r catalysis, provided th bas -pairing of helix P1 is maintained.

- 21 -

Third, the exon sequences flanking the 3' splice site must be chosen, and adjustments made in the 5' section of the ribozyme, if necessary, to allow the formation of a stable P10 helix. While the P10 helix may be dispensed with if necessary, its presence enhances splicing and preferred embodiments of the ribozyme of the invention retain the P10 helix (Suh, E.R. et al., "Base Pairing Between The 3' Exon And An Internal Guide Sequence Increases 3' Splice Site Specificity in the *Tetrahymena* Self-Splicing rRNA Intron," *Mol. Cell. Biol.* 10:2960-2965 (1990)). The helices P1 and P10 overlap along the *T. thermophila* intron IGS, and the 2nd and 3rd residues following both the 5' and 3' splice sites are complementary to the same residues in the IGS (Figure 2). While there may be some advantage in following this, many natural group I introns do not share this constraint, so the choice of 3' exon sequences may be determined primarily by experimental considerations. Such considerations reflect the wide flexibility in choice of splice sites. For example, if it is desired to join two sequences at a given point, the sequence at such point cannot be mutated or otherwise altered by the trans-splicing event. Either P1 or P10 can be made shorter if the overlapping sequences don't otherwise accommodate for the desired splice site.

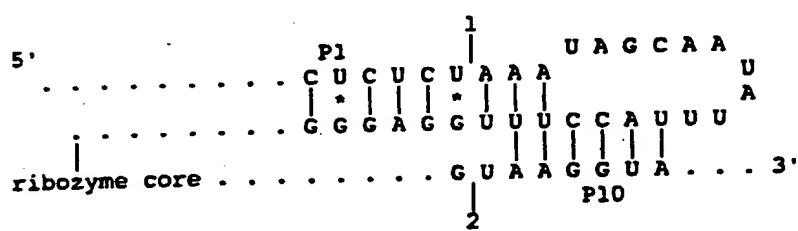
The sequence requirements for 3' splice-site selection appear to lie mainly within the structure of the intron (the ribozyme) itself, including helix P9.0 and the adjoining 3' guanosine residue which delineates the 3' intron boundary. P9.0 is wholly contained within the intron sequences and helps define the adjacent 3' splice site. For the trans-splicing design, the P9.0 helix and the rest of the functional RNA elements within the intron are not altered. The structural characteristics of the P9.0 helix are known (Michel, F.

- 22 -

t al., "The Guanosine Binding Site of the Tetrahymena Ribozyme," Natur 342:391-395 (1989)). However, flanking sequences within the 3' exon are required for the formation of helix P10 and efficient splicing, as shown by mutational analysis.

Fourth, a region of complementary sequence is placed at the 5' terminus of the trans-splicing ribozyme in order to increase its affinity and specificity for the target RNA. As shown herein, an arbitrary length of around 40 residues has been used. Other lengths may be used provided they are not detrimental to the desired effect.

For example, starting with the *T. thermophila* self-splicing intron (diagrammed below):



(The "1" and "2" in the above diagram (and in other ribozyme diagrams throughout the application) note the first and second splice sites, respectively.)

(1) a "5'" site is chosen adjacent to a uracil residue within a chosen target RNA. The sequences involved in complementarity do not immediately abut sequences involved in P1 helix formation but are separated, for example, by five nucleotides also involved in P10 formation;

(2) sequences complementary to the chosen RNA are fused to the 5' portion of the self-splicing Group I intron. Base-pairing between ribozyme and target RNA allow formation of the helix P1;

(3) the chosen "3' exon" sequences are fused to the 3' portion of the ribozyme, maintaining the conserved helix P10; and

- 23 -

(4) to increase affinity for the target RNA, if desired, a section of extended sequence complementary is fused to the 5' portion of the ribozyme to allow the formation of 30-40 base-pairs.

5 The alignment of the resulting trans-splicing ribozyme with its target RNA may be diagrammed as shown immediately below. The target RNA sequence represents the top line. The ribozyme sequence is aligned below it, a continuous sequence wrapping around the lower two lines wherein the hybridization of the nucleotides at the 5' and 3' ends and P1 and P10 of the ribozyme may 10 be seen.

## Alignment of the Ribozyme of the Invention with a Target RNA

- 25 -

5 According to the invention, trans-splicing ribozymes can be designed that will trans-splice essentially any RNA sequence onto any RNA target. It is not necessary that the target contain an intron sequence or that the ribozyme be an intron in the target sequence. For example, a strategy for such design may include (1) the identification of the desired target RNA (2) cloning and/or sequencing of the desired target RNA or portion thereof (3) selection of a desired coding sequence to trans-splice into the target RNA, (4) the construction of a ribozyme of the invention capable of hybridizing to such target using the guidelines herein and (5) confirmation that the ribozyme of the invention will utilize the target as a substrate for the specific trans-splicing reaction that is desired and (6) the insertion of the ribozyme into the desired host cell.

10

15

20 Choice of a target RNA will reflect the desired purpose of the trans-splicing reaction. If the purpose of the reaction is to inactivate a specific RNA, then such RNA must be trans-spliced at a position that destroys all functional peptide domains encoded by such RNA and at a position that does not result in continued expression of the undesired genetic sequences. If more than one allele of the gene encoding such RNA exists, the ribozyme should preferably be designed to inactivate the target RNA at a site common to all expressed forms. Alternatively, more than one ribozyme may be provided to the cell, each designed to inactivate a specific allelic form of the target RNA.

25

30

When only inactivation of the target RNA is desired, and not the expression of a new, desired RNA sequence, it is not necessary that the foreign RNA donated by the ribozyme provide a sequence capable of

- 26 -

being translated by the host cell, and a sequence containing translational stop codons may be used as a truncated intron, for example, the intron ribozyme truncated at the *Sca*I site.

5 If the purpose of the trans-splicing reaction is to provide a genetic trait to a host cell, then the choice of target RNA will reflect the desired expression pattern of the genetic trait. If it is desired that the genetic trait be continuously expressed by the host, then the target RNA should also be continuously expressed. If it is desired that the genetic trait be selectively expressed only under a desired growth, hormonal, or environmental condition, then the target RNA should also be selectively expressed under such conditions.

10 It is not necessary that expression of the ribozyme itself be selectively limited to a desired growth, hormonal, or environmental condition if the substrate for such ribozyme is not otherwise present in the host as the ribozyme itself is not translated by the host. Thus, sequences encoded by the RNA donated by the ribozyme of the invention are not translated in a host until the trans-splicing event occurs and such event may be controlled by the expression of the ribozyme substrate in the host.

15 If desired, expression of the ribozyme may be engineered to occur in response to the same factors that induce expression of a regulated target, or, expression of the ribozyme may be engineered to provide an additional level of regulation so as to limit the occurrence of the trans-splicing event to those conditions under which both the ribozyme and target are selectively induced in the cell, but by different factors, the combination of those factors being the undesired event. Such regulation would allow the host cell to express the ribozyme's target under those

- 27 -

conditions in which the ribozyme itself was not co-expressed.

5 The sequence of the ribozyme domain that hybridizes to the target RNA is determined by the sequence of the target RNA. The sequence of the target RNA is determined after cloning sequences encoding such RNA or after sequencing a peptide encoded by such target and deducing an RNA sequence that would encode such a peptide. Cloning techniques known in the art 10 may be used for the cloning of a sequence encoding a target RNA.

15 The selection of a desired sequence to be trans-spliced into the target RNA (herein termed the "trans-spliced sequence") will reflect the purpose of the trans-splicing. If a trans-splicing event is desired that does not result in the expression of a new genetic sequence, then the trans-spliced sequence need not encode a translatable protein sequence. If a trans-splicing event is desired that does result in the 20 expression of a new genetic sequence, and especially a new peptide or protein sequence, then the trans-spliced sequence may further provide translational stop codons, and other information necessary for the correct translational processing of the RNA in the host cell. If a specific protein product is desired as a result of the trans-splicing event then it would be necessary to 25 maintain the amino acid reading frame in the resulting fusion.

30 The identification and confirmation of the specificity of a ribozyme of the invention is made by testing a putative ribozyme's ability to catalyze the desired trans-splicing reaction only in the presence of the desired target sequence. The trans-splicing reaction should not occur if the only RNA sequences 35 present are non-target sequences to which such ribozyme should not be responsive (or less responsive). Such

- 28 -

5

characterization may be performed with the assistance of a marker such that corr ct (or incorr ct) rib zyme activity may be more easily monitored. In most cases it is sufficient to test the ribozyme against its intended target *in vitro* and then transform a host cell with it for study of its *in vivo* effects.

10

When it is desired to eliminate a host's RNA, such elimination should be as complete as possible. When it is desired to provide a new genetic sequence to a host cell, the trans-splicing reaction of the invention need not be complete. It is an advantage of the invention that, depending upon the biological activity of the peptide that is translated from such genetic sequence, the trans-splicing event may in fact be quite inefficient, as long as sufficient trans-splicing occurs to provide sufficient mRNA and thus encoded polypeptide to the host for the desired purpose.

15

Transcription of the ribozyme of the invention in a host cell occurs after introduction of the ribozyme gene into the host cell. If the stable retention of the ribozyme by the host cell is not desired, such ribozyme may be chemically or enzymatically synthesized and provided to the host cell by mechanical methods, such as microinjection, liposome-mediated transfection, electroporation, or calcium phosphate precipitation. Alternatively, when stable retention of the gene encoding the ribozyme is desired, such retention may be achieved by stably inserting at least one DNA copy of the ribozyme into the host's chromosome, or by providing a DNA copy of the ribozyme on a plasmid that is stably retained by the host cell.

20

25

30

35

Preferably the ribozyme of the invention is inserted into the host's chromosome as part of an expression cassette, such cassette providing transcriptional regulatory elements that will control the transcription of the ribozyme in the host cell.

- 29 -

Such elements may include, but not necessarily be limited to, a promoter element, an enhancer or UAS element, and a transcriptional terminator signal. Polyadenylation is not necessary as the ribozyme is not translated. However, such polyadenylation signals may be provided in connection with the sequence encoding the element to be trans-spliced.

Expression of a ribozyme whose coding sequence has been stably inserted into a host's chromosome is controlled by the promoter sequence that is operably linked to the ribozyme coding sequences. The promoter that directs expression of the ribozyme may be any promoter functional in the host cell, prokaryotic promoters being desired in prokaryotic cells and eukaryotic promoters in eukaryotic cells. A promoter is composed of discrete modules that direct the transcriptional activation and/or repression of the promoter in the host cell. Such modules may be mixed and matched in the ribozyme's promoter so as to provide for the proper expression of the ribozyme in the host. A eukaryotic promoter may be any promoter functional in eukaryotic cells, and especially may be any of an RNA polymerase I, II or III specificity. If it is desired to express the ribozyme in a wide variety of eukaryotic host cells, a promoter functional in most eukaryotic host cells should be selected, such as a rRNA or a tRNA promoter, or the promoter for a widely expressed mRNA such as the promoter for an actin gene, or a glycolytic gene. If it is desired to express the ribozyme only in a certain cell or tissue type, a cell-specific (or tissue-specific) promoter elements functional only in that cell or tissue type should be selected.

The trans-splicing reaction is chemically the same

- 30 -

whether it is performed *in vitro* or *in vivo*. However, *in vivo*, since cofactors are usually already present in the host cell, the presence of the target and the ribozyme will suffice to result in trans-splicing.

5 The trans-splicing ribozymes and methods of the invention are useful in producing a gene activity useful for the genetic modification, and/or cell death, of targeted cells. For example, the trans-splicing reaction of the invention is useful to introduce a protein with toxic properties into a desired cell. The 10 susceptibility of cells will be determined by the choice of the target RNA and the regulatory controls that dictate expression of the ribozyme. For example, a ribozyme that transposes an RNA sequence encoding a 15 toxic protein may be engineered so that expression of the ribozyme will depend upon the characteristics of an operably-linked promoter. In a highly preferred embodiment, diphtheria toxin peptide A is encoded by that part of the ribozyme that is transposed into a 20 desired target in the host. Conditional expression of the ribozyme and diphtheria toxin peptide A chain results in the death of the host cell. Other useful 25 peptide toxins include ricin, exotoxin A, and herpes thymidine kinase (Evans, G.A., *Genes & Dev.* 3:259-263 (1989)). In addition, various lytic enzymes have the potential for disrupting cellular metabolism. For 30 example, a fungal ribonuclease may be used to cause male sterility in plants (Mariani, C. et al., *Nature* 347:737-741 (1990)). Particular tissues might be destroyed due to limited expression of the target RNA. Further, if a viral RNA is used as target, new forms of 35 virus resistance, or therapies may be engineered.

A binary system for control of tissue-specific gene expression and/or for ectopic ablation may be designed using the ribozymes of the invention. For example, lines of *Drosophila* that express the yeast

- 31 -

transcription activator GAL4 in a tissue and spatial-specific pattern using P-element enhancer-trap vectors may be used. Any transcriptional activator may be used in place of GAL4 and the invention is not intended to be limited to GAL4. A gene encoding a fusion ribozyme that is capable of trans-splicing the DTA sequence may be placed under the control of the GAL4-UAS promoter and inserted into *Drosophila* in a genetically stable manner. Such ribozymes will not be expressed in *Drosophila* in the absence of GAL4. Accordingly, crossing *Drosophila* hosts genetically carrying this ribozyme construct with *Drosophila* hosts that express GAL4 in a tissue-specific manner result in progeny that, when GAL4 expression is induced, exhibit a pattern of cell death similar to the pattern of GAL4 expression.

In addition, by targetting the ribozyme to trans-splice with the GAL4 mRNA, the splicing activity of the ribozyme inactivates GAL4 expression and ribozyme expression may be self-regulated.

20

#### Pre-ribozymes

25

30

35

A trans-splicing ribozyme, as described above, consists of three fused sequence elements - a 5' "anti-sense" region which is complementary to the target RNA, the catalytic region which is based on a self-splicing Group I intron, and 3' "exon" sequences. The 5' region can base pair with the chosen target RNA, to bring it into proximity with the catalytic sequences of the Group I intron. The structure of the Group I intron provides a chemical environment suitable to catalyze the precise splicing of the target RNA with the 3' "exon" sequences. However, in the absence of the appropriate target RNA, the ribozyme sequences can still catalyze scission at the 3' "exon" junction (similar hydrolysis is seen for Group I self-splicing

- 32 -

intons (Zaug et al., *Science* 231:470-475 (1986)), and may be able to catalyze illegitimate splicing events through transient base-pairing of the ribozyme with heterologous RNA sequences (which may include their own). Such side-reactions and illegitimate splicing events are unwanted, and may be deleterious. For example, if trans-splicing is to be used for conditional delivery of a toxin *in vivo*, illegitimate trans-splicing might result in unexpected expression of the toxic activity. Spontaneous cleavage at the 3' "exon" junction would lower the efficiency of trans-splicing.

To help avoid these problems, "pro-ribozyme" forms of the trans-splicing RNAs have been constructed wherein for example, helix P8 is disrupted. The pro-ribozymes are constructed to contain extra self-complementary sequences which cause the catalytic center of the ribozyme to be mis-folded. The pro-ribozymes are inactive in the absence of the intended target RNA; active forms are only formed after base-pairing of the ribozyme and target RNAs with consequent displacement of the interfering secondary structure within the ribozyme. Pro-ribozymes are intended to be catalytically inert species in the absence of the target RNA, to eliminate unwanted self-cleavage, self-splicing and illegitimate trans-splicing reactions *in vitro* and *in vivo* (Figure 12).

The pro-ribozymes described here are conformationally disrupted and therefore inactive forms of the trans-splicing activities. Thus the pro-ribozymes possess little self-cleavage activity. They are only re-activated by specific interaction with the target RNA, and thus are substrate-activated ribozymes which are less likely to catalyze trans-splicing to an unintended target RNA. Trans-splicing ribozymes are intended to be used for the delivery of new gene

- 33 -

activities in vivo, and any reduction in the extent of unwanted side reactions or illegitimate splicing is desirable, and may be necessary.

5

While the disruption of helix P8 has been exemplified here for the trans-splicing pro-ribozymes, other helices which are required for catalytic activity could also have been used.

10

The same approach, of disrupting the conformation of a catalytically important structure in such a way that only base-pairing with the intended substrate RNA will allow the formation of an active ribozyme, could be applied to other ribozyme designs. For example, the loop sequence of a "hammerhead" type endoribonuclease (Haseloff et al., *Nature* 334:585-591 (1988)) could be extended and made complementary to one of the "anti-sense" arms of the ribozyme - similar to the above modification of helix P8. Endoribonuclease activity would only be exhibited after base-pairing with the chosen target RNA, displacement of the disrupting secondary structure, and reformation of the stem-loop structure required for catalysis. This would effectively increase the specificity of the ribozyme of its target.

20

In addition, the activation of a pro-ribozyme need not rely on base-pairing with the substrate itself. Instead, a chosen third RNA or ssDNA or even protein might be required for activity. An additional base-pairing or RNA-protein interaction would be required for the formation of an active ribozyme complex. The availability of such additional components would determine ribozyme activity, and could be used to alter ribozyme selectivity.

30

The ribozyme or pro-ribozyme of the invention may be introduced into any host cell, prokaryotic or eukaryotic and especially into a plant or mammalian host cell, and especially a human cell, either in

35

- 34 -

culture or in vivo, using techniques known in the art appropriate to such hosts. The ribozymes of the invention may also be engineered to destroy viruses. In one embodiment, the ribozyme or pro-ribozyme of the invention is provided in a genetically stable manner to a host cell prior to a viral attack. Infection by the appropriate virus, or expression of the latent virus in such host cell, (resulting in the appearance of the ribozyme's or pro-ribozyme target RNA in the host cell), would stimulate the catalytic activity of the ribozyme and destruction of the viral RNA target and/or production of a toxin via trans-splicing resulting in death of the virus infected cells. In another embodiment, the ribozyme or pro-ribozyme may be engineered and packaged into the virus itself. Such embodiments would be especially useful in the design of viruses for investigative purposes, wherein the ribozyme or pro-ribozyme may be designed to destroy the function of a specific viral RNA and thus allow the study of viral function in the absence of such RNA. Viruses carrying ribozymes may also be used as carriers to transfect host cells with a desired ribozyme or pro-ribozyme activity.

Male or female sterility may be engineered in agronomically important species using the ribozymes or pro-ribozymes of the invention. For example, male sterility in tobacco may be engineered by targetting TA29 or TA13 mRNA (tobacco anther-specific genes; Seurinck, J. et al., Nucl. Acids Res. 18:3403 (1990) with a ribozyme or pro-ribozyme of the invention that trans-splices the DTA 3' exon into those targets.

The form of crop plants may be manipulated by selective destruction or modification of tissues using the ribozymes or pro-ribozymes of the invention. For example, seedless fruits may be made by targetting the storage protein mRNA with a ribozyme or pro-

- 35 -

ribozyme of the invention that trans-splices the DTA 3' exon into the target.

Transgenic plants may be protected against infection by expression of virus-specific ribozymes or pro-ribozyme to kill infected cells. This would be an artificial form the "hypersensitive response." For example, cucumber mosaic virus coat protein mRNA may be targeted with a ribozyme or pro-ribozyme of the invention that trans-splices the DTA 3' exon into the target.

Populations of micro-organisms may be made resistant to specific pathogens by introduction of trans-splicing ribozymes or pro-ribozymes. For example, cheese-making bacteria may be made resistant to phage infection by targetting the phage RNA with a bacterial toxin gene or lytic enzyme encoded by the 3' exon provided by the ribozyme or pro-ribozyme of the invention, for example, which would interfere with phage replication by causing premature lysis after phage infection.

Virus pathogens could be constructed to deliver toxic activities via trans-splicing. In this way, specific cell types could be targeted for ablation, such as for cancer or viral therapy. For example, HIV mRNA may be targeted by a ribozyme or pro-ribozyme of the invention that carries the DTA 3' exon, for either virus or liposome delivery.

The examples below are for illustrative purposes only and are not deemed to limit the scope of the invention.

- 36 -

## EXAMPLES

### Example 1

#### Construction and Characterization of a CAT-LacZ Trans-Splicing Ribozyme

5

10

#### I. PCR Amplification and Cloning of the Ribozyme of the Invention

15

20

25

30

35

40

Following the guidelines outlined above, a trans-splicing fusion ribozyme was designed that will splice a portion of the amino-terminal coding sequence of *E. coli*  $\beta$ -galactosidase (LacZ) mRNA to a site in the chloramphenicol acetyl transferase (CAT) mRNA (Figure 3). The sections of new sequence flanking the *T. thermophila* ribozyme core and the 3' exon were synthesized as oligonucleotides. The intact ribozyme sequence was then assembled by successive polymerase chain reactions, using the synthetic adaptor oligonucleotides as primers with ribozyme and  $\beta$ -galactosidase DNA templates (while there are other methods available, this method is most convenient).

For the construction of a ribozyme capable of splicing  $\beta$ -galactosidase (LacZ)  $\alpha$ -peptide coding sequence to a site in the 5' coding sequence of the chloramphenicol acetyl transferase (CAT), three oligonucleotides were synthesized.

##### Oligonucleotide 1

5'-GGCCA AGCTT CTTTA CGATG CCATT GGGAT ATATC AACGG  
TGGTA TAAAC CCGTG GTTTT TAAAA GTTAT CAGGC ATGCA CC-3'  
[SEQ ID NO. 2]

##### Oligonucleotide 2

5'-GATTA GTTTT GGAGT ACTCG TACGG ATTCA CGGCC GTCGT  
TTTAC AA-3' [SEQ ID NO. 3]

- 37 -

Oligonucleotide 3

5'-GGCCG AATTC TTACA ATTTC CATTC AGGCT GCGCA ACTGT TGG-  
3' [SEQ ID NO. 4]

5 Oligonucleotides 2 and 3 (200 pmoles each) were combined with 0.1  $\mu$ g *Pvu*II-cut pGEM4 DNA (which contained the LacZ  $\alpha$ -peptide sequence), and subjected to PCR amplification in a volume of 100  $\mu$ l containing:

10 50 mM KCl,  
10 mM Tris-HCl pH 8.3,  
1.5 mM MgCl<sub>2</sub>,  
0.4 mM dNTPs,  
15 0.1% gelatin, and  
5 U *Taq*I DNA polymerase,

20 and incubated for 30 cycles, 1 min @ 94°C, 2 mins @ 50°C, 2 mins @ 72°C.

Plasmid pGEM4 is commercially available from Promega Corporation, Madison WI, USA.

25 The amplified product of 210 base-pairs was purified using low-gelling temperature agarose electrophoresis, and was used as primer in a second round of PCR amplification.

30 Following the second round of PCR amplification, 2.0  $\mu$ g of 210 base-pair amplified product, 200 pmoles oligonucleotide 1 and 0.1  $\mu$ g 450 base-pair fragment containing the *T. thermophila* IVS were mixed and subjected to PCR amplification using the conditions shown above. The resulting 660 base-pair product was digested with the restriction endonucleases EcoRI and 35 *Hind*III, and cloned into the plasmid vector pGEM4. The complete sequence of the CAT-LacZ  $\alpha$ -peptide ribozyme DNA sequence is presented as SEQ ID NO. 5 and Figure 3B.

40 The cloning vector containing the cloned sequences was transformed into, and propagated in, the bacterial host XL1/Blue (Strategene, La Jolla, California), using techniques known in the art (Maniatis, Molecular

- 38 -

Cloning, A Laboratory Guide, 2nd edition, 1989, Cold Spring Harbor Laboratory, Publishers). However, any bacterial host capable of stably maintaining the vector may be used, for example the JM109.

5 The plasmid may be extracted from the host cell for further analysis using techniques commonly known in the art (Maniatis, Molecular Cloning, A Laboratory Guide, 2nd edition, 1989, Cold Spring Harbor Laboratory, Publishers).

10 **II. In vitro Transcription of Cloned Ribozyme and Target RNAs**

15 Using standard procedures, cloned sequences were purified from the bacterial host and the plasmid linearized using a restriction endonuclease that does not cut the ribozyme sequence, (for example, EcoRI), and transcribed using T7 RNA polymerase in a volume of 20 100  $\mu$ l, containing:

25 5  $\mu$ g linearized plasmid DNA,  
40 mM Tris-HC pH 7.5,  
6 mM MgCl<sub>2</sub>,  
2 mM spermidine,  
10mM NaCl,  
10mM DTT,  
30 1mM NTPs (containing 20  $\mu$ Ci [ $\alpha$ -<sup>32</sup>P]UTP, if labelled RNA transcripts were desired),  
100 U RNasin, and  
50 U T7 RNA polymerase,

and the reaction was incubated at 37°C for 2 hours.

35 RNA transcripts were purified by 5% polyacrylamide gel electrophoresis before use (TBE, 7M urea gel). RNAs containing active *T. thermophila* IVA sequences undergo some spontaneous scission at the 3' intron-exon junction during transcription. Fragments are removed by electrophoretic purification for clarity of analysis 40 during subsequent trans-splicing assays.

- 39 -

### III. In Vitro Trans-splicing Reaction Conditions

Target and/or trans-splicing ribozymes are incubated under the following conditions:

5           0.1-0.5 µg RNA component (amount depends on type of experiment, usually ribozyme in 5-fold excess of target),  
          30 mM Tris-HCl pH 7.5,  
          100 mM NaCl,  
10           2mM GTP,  
          5 mM MgCl<sub>2</sub>,  
          in a volume of 5 µl at 42°C, 60 mins.

15           The reaction is diluted with 95 µl 0.1 mM Na<sub>2</sub>EDTA, 200 mM NaCl, and ethanol precipitated. The RNAs are then analysed on 5% polyacrylamide gels containing TBE buffer, 7M urea and 25% formamide, and autoradiographed.

### IV. Assay of Endonucleolytic Activity

20           After base-pairing of the ribozyme and target, the first step in trans-splicing is the guanosine mediated cleavage of the target RNA at the intended 5' splice site. Annealing and trans-splicing may be performed in a buffer such as 30 mM Tris-HCl, pH 7.5, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, 2 mM GTP at 42°C. As the 3' splice site is dispensable for this reaction, truncated trans-splicing ribozymes should behave as highly-specific endoribonucleases. To test this activity, shortened in vitro transcripts of the CAT-LacZ α-peptide trans-splicing ribozyme described above (SEQ ID NO. 5 and Figure 3) were incubated with CAT mRNA sequences. The CAT-LacZ ribozyme cassette is on a HindIII-EcoRI fragment. The ScaI cleavage site marks a position 5 bases upstream of the 3' splice site. The ribozyme specifically cleaved the target RNA at the expected single site to produce the expected size fragments.

- 40 -

**V. The Trans-splicing Reaction**

To confirm the ability of the CAT-LacZ  $\alpha$ -peptide 5  
ribozyme to catalyze the ligation of 3' exon sequences  
at the 5' splice site, various forms were incubated  
with radiolabelled CAT RNA. Ribozyme transcripts were  
synthesized from DNA templates which had been 3'  
truncated at one of several positions, ranging from the  
end of the ribozyme core through the exon sequence.  
10 Incubation with labelled CAT led to the formation of  
the expected spliced products, which differed in length  
depending on the extent of 3' exon sequence.

In addition, a certain proportion of the CAT-LacZ  
apeptide ribozyme molecules underwent spontaneous  
15 cleavage at the 3' splice site during in vitro  
transcription, similar to the intact *T. thermophila*  
intron. These cleaved forms, terminated at the  
guanosine residue adjacent the 3' splice site, were  
also incubated with CAT RNA. In this case, the ribozyme  
itself is ligated to a 3' portion of the CAT RNA, to  
produce a product of about 550 nucleotides in size.  
20 This reaction is similar to the self-circularization of  
the intact intron, and the same ligation product is  
found in the other trans-splicing reactions.

**VI. Accuracy of the Trans-splicing**

The products from a CAT-LacZ  $\alpha$ -peptide trans-  
splicing reaction were reverse-transcribed, and  
30 amplified by polymerase chain reaction using two  
oligonucleotides complementary to sequences on either  
side of the predicted splice sites. Amplified  
sequences were cloned and sequenced. Individual  
recombinants showed no variation from the expected  
35 sequences of the spliced products. As found in studies

- 41 -

with the intact intron, splicing appears to be highly accurate.

Accordingly, the studies above show that a trans-splicing ribozyme designed according to the guidelines of the invention is capable of accurate, effective trans-splicing <sup>5</sup> *in vitro*.

Example 2

Design of a Trans-Splicing Ribozyme that Provides  
Plant Virus Resistance

Cucumber mosaic virus (CMV) is a pandemic virus with a large number of known strains. Nine sequence strains are shown in the region of the start of their <sup>15</sup> coat protein cistron encoded in RNA 3 and the subgenomic mRNA 4 (SEQ ID NOS. 7-25; Figures 4(A) and 5). Two sites have been chosen which are conserved in sequence and downstream from the AUG start codon of the coat protein. Oligonucleotides for the construction of <sup>20</sup> ribozymes capable of trans-splicing the ile-mutant form of DTA into the CMV coat protein mRNA are shown in Figure 4B and is discussed below.

The trans-splicing ribozymes shown in Figure 4C and D are targetted to the CMV virus sequences shown in <sup>25</sup> Figure 4B and will result not only in the cleavage of the CMV RNA molecules but in the expression of diphtheria toxin A-chain in the infected cell. The trans-splicing cassettes shown in Figure 4 may be transformed into any CMV-susceptible plant species <sup>30</sup> using techniques known in the art, and transgenic progeny challenged by CMV infection. The design of the ribozyme is such that virus infection is necessary to initiate toxin production via RNA trans-splicing <sup>35</sup> because the ribozyme itself is not translated. The localized death of the infected cells that results from expression of the toxin could limit replication and

- 42 -

spread of the virus within the plant giving an artificial hypersensitive response.

5

Example 3

Construction and Characterization of a Gal4-Diphtheria Toxin A Chain Trans-Splicing Ribozyme

10 According to the invention and the methods described in Example 1, a fusion ribozyme has been designed that is a Gal4-Diphtheria toxin A chain trans-splicing ribozyme (Figure 6). The sequence of this ribozyme is shown as SEQ ID NO. 6. The GAL4-DTA ribozyme cassette is a *Sca*I-*Xho*I fragment. The *Sca*I site marks a position 5 bases upstream of the 3' splice site. This ribozyme is capable of splicing the coding sequence for the A chain of the diphtheria toxin to a site in the 5' region of the GAL4 mRNA. This trans-splicing activity is active both *in vitro* (as above) and *in vivo* (below). The major criteria for successful design of the GAL4-DTA ribozyme, and any trans-splicing ribozyme that trans-splices a sequence encoding a toxic product, are not only the efficient and precise catalysis of trans-splicing, but also that expression of the toxic product, for example, DTA does not occur in the absence of trans-splicing.

15

20

25

30

The catalytic portion of the ribozyme is constructed according to the design outlined above, and 5' and 3' splice sites chosen within the 5' coding regions of GAL4 and DTA, respectively. The 3' exon sequence corresponds to that of a DTA gene already used for expression in eukaryotes, except for the removal of the first AUG codon and several proximal amino acids. The original *C. diphtheriae* form of DTA also differs in this 5' region, utilizing a CUG codon for translation

35

- 43 -

initiation. The original DTA sequence also contains a signal peptide leader sequence which is absent.

These ribozyme molecules can undergo spontaneous scission at the 3' splice site. Given the extreme toxicity of DTA, it is important that any liberated 3' exon sequences not give rise to toxic translation products. The 3' exon contained an in-frame methionine at position 13, which could conceivably give rise to a truncated but toxic polypeptide. To eliminate this possibility, the wild-type sequence (Rz-DTA<sub>met</sub>) was altered from methionine at this position to isoleucine (Rz-DTA<sub>ile</sub>) or leucine (Rz-DTA<sub>leu</sub>) in two separate ribozyme constructions (Figure 6).

15

#### Example 4

#### In Vivo Activity of the Ribozymes of the Invention

##### I. Introduction

20 The *in vivo* activity of a ribozyme designed according to the guidelines provided herein, and the ability of such a ribozyme to deliver new gene activities to host cells, was demonstrated using the Gal4-Diphtheria toxin A chain trans-splicing ribozyme described (Example 3 and in Figure 6) to deliver the highly toxic diphtheria toxin A product to a host cell. In this system, *Drosophila* was the chosen host and it was desired to control expression of the ribozyme of the invention in a tissue-specific manner within the *Drosophila* host.

25 Diphtheria toxin is secreted by *Corynebacterium diphtheriae* lysogenic for B phage. The toxin is produced as a single polypeptide which undergoes proteolysis to produce A and B chains. The A chain (DTA) contains a potent ADP ribosylase activity which is specific for the eukaryote translation elongation factor EF-2. The presence of even a few molecules of

35

- 44 -

5 this nzyme is enough to cause cessation of translation and eventual death in a variety of eukaryote cells. The B chain allows intracellular delivery by attachment of the toxin to cell surface receptors by binding mannose residues, is endocytosed and enters the cytoplasm by vesicular fusion.

10 In the absence of the B-chain, the A-chain is much less toxic when present extracellularly. This property, and its extreme toxicity, have suggested its use for ectopic ablation experiments. For example, sequences encoding DTA have been expressed in transgenic mice, using an opsin promoter to drive expression in developing eyes. The resulting mice are blind, with deformed eyes (Breitman, M.L., *Science* 238:1563-1565 (1987)). In other studies, ablation of the mouse pancreas was performed (Palmeter, R.D. et al., *Cell* 50:435-443 (1987)) and Wert, S.E. et al., *Am. Rev. Respir. Dis.* 141 (no. 4, part 2):A695 (1990) described ablation of alveolar cells by use of a chimeric gene consisting of the promoter and 5' flanking sequence of the human surfactant protein C gene (expressed in type II alveolar cells) and the DTA gene.

20 25 30 However, using this type of approach, it is not possible to maintain or propagate transformed organisms which might have more severe, or lethal phenotypes. In addition, transformation of certain species, such as *Drosophila*, with intact DTA sequences has not been reported to date. Leaky expression of the DTA gene during such transformations leads to immediate death.

## II. The Drosophila System

35 A general method for targeting gene expression in *Drosophila* has been developed. First, the system allows the rapid generation of individual strains in

- 45 -

which ectopic gene expression can be directed to different tissues or cell types: the enhancer detector technique is utilized (O'Kane, C.J. and Gehring, W.J., *Proc. Natl. Acad. Sci. USA*:9123-9127 (1987); Bellen et al., *Genes and Development* 3:1288-1300 (1989); Bier et al., *Genes and Development* 3:1273-1287 (1989)) to express a transcriptional activator protein in a wide variety of patterns in embryos, in larvae and in adults. Second, the method separates the activator from its target gene in distinct lines, to ensure that the individual parent lines are viable: in one line the activator protein is present but has no target gene to activate, in the second line the target gene is silent. When the two lines are crossed, the target gene is turned on only in the progeny of the cross, allowing dominant phenotypes (including lethality) to be conveniently studied.

To ectopically express only the gene of interest, a transcriptional activator that has no endogenous targets in flies is required. An activator from yeast, Gal4, can activate transcription in flies but only from promoters that bear Gal4 binding sites (Fischer et al., *Nature* 332:853-865 (1988)). To target gene expression, Gal4 is restricted to particular cells in two ways: either Gal4 transcription is driven by characterized fly promoters, or an enhancerless Gal4 gene is randomly integrated in the *Drosophila* genome, bringing it under the control of a diverse array of genomic enhancers. To assay transactivation by Gal4, flies that express Gal4 are crossed to those bearing a lacZ gene whose transcription is driven by Gal4 binding sites (Fischer et al., *Nature* 332:853-865 (1988)).  $\beta$ -galactosidase is expressed only in those cells in which Gal4 is first expressed. Tissue- and cell-specific transactivation of lacZ has been demonstrated in strains in which Gal4

- 46 -

is expressed and in which a variety of patterns are established.

With this system, it is now possible: 1) to place Gal4 binding sites upstream of any coding sequence; 2) to activate that gene only within cells where Gal4 is expressed and 3) to observe the effect of this aberrant expression on development. In cases where ectopic expression is lethal, this method allows the two parent lines (one expressing Gal4, the other carrying a silent gene bearing Gal4 binding sites in its promoter) to be stably propagated. Phenotypes can then be studied in the progeny of a cross.

### III. Vectors

The vectors utilized as starting materials in these studies include:

1) pGATB and pGATN (figure 9): These vectors are used for cloning promoters and enhancers upstream of a promoterless Gal4 gene.

Vectors were constructed in which either a unique NotI or BamHI site is inserted upstream of the Gal4 coding region. Once a promoter has been linked to the Gal4 coding sequence, the gene can be excised from the pHsREM vector backbone (Knipple and Marsella-Herrick, *Nucl. Acids Res.* 16:7748 (1988)) and moved into a P-element vector. The Rh2 promoter has been cloned (Mismer et al., *Genetics* 120:173-180 (1988)) into this vector and flies have been generated in which Gal4 is expressed only in the ocelli.

2) pGawB: This is a Gal4 vector for use in enhancer detection.

An enhancerless Gal4 gene was subcloned into the vector plwB (Wilson et al., *Genes and Development* 3:1301-1313 (1989)) to create pGawB. plwB was first digested with HindIII to remove the *lacZ* gene and the N-terminus of the P-transposase gene. These were

- 47 -

replaced with the entire Gal4 coding region behind the TATA box of the P-transposase gene.

3) pUAST (Figure 10): This plasmid was used for cloning coding sequences downstream of the Gal UAS.

A vector into which genes can be subcloned behind the Gal4 UAS (Upstream Activation Sequence) was constructed in the P-element vector, pCaSpeR3 (C. Thummel, Univ. of Utah Medical Center, Salt Lake City, Utah, personal communication). Five Gal4 binding sites were inserted, followed by the hsp70 TATA box and transcriptional start, a polylinker, and the SV40 intron and polyadenylation site. Unique sites into which genes, or cDNAs, can be inserted include: EcoRI, BglIII, NotI, XhoI, KpnI and XbaI.

#### IV. Drosophila Strains

The genetic techniques described herein used to characterize the strains of *Drosophila* utilized in these studies are well known in the art ("Genetic Variations of *Drosophila melanogaster*," D. Lindsley and E.H. Grell, eds.).

The P-element transposons are mobilized using the "jumpstarter" strain that carries  $\Delta 2-3$ , a defective P-element on the third chromosome that expresses high levels of a constitutively active transposase (Robertson et al., *Genetics* 118:451-470 (1988)). The three stocks currently used to generate and map the insertion lines were deposited in the *Drosophila* Stock Center, Indiana University Department of Biology, Jordan Hall A 503, Bloomington, Indiana 47405:

1: y w; +/+; Sb P[ry<sup>+</sup>,  $\Delta 2-3$ ]/TM6, Ubx

2: w; +/+; TM3, Sb/CxD (deposit no. 3665)

3: w; CyO/Sco; +/+ (deposit no. 3666)

where the genetic characteristics of the three chromosomes are separated by semicolons. Thus, for example, in strain 1, the first chromosome (the X

- 48 -

chromosome) is homozygous for yellow and white ("y w"), the second chromosome is wild-type ("+/+"), and the third chromosome carries the stubble gene ("Sb"), and the P element transposon rosy gene ("ry<sup>+</sup>") and  $\Delta 2-3$ , while the second third chromosome carries balancer inversions ("TM6, Ubx").

v. Strategy for Generating Gal4 Expression Patterns

10 A. Scheme used to isolate transformants

Constructs are injected into embryos derived from the stock;

♀♀ y w/y w ;  $\Delta 2-3$ , Sb/TM6, Ubx X ♂♂ y w/Y ;  $\Delta 2-3$ , Sb/TM6, Ubx

FO; Establish single lines

♀ y w/y w ;  $\Delta 2-3$ , Sb/TM6, Ubx X ♂ y w/Y; +/+

or

♂ y w/Y ;  $\Delta 2-3$ , Sb/TM6, Ubx X ♀ y w/y w; +/+

F1; Select [w<sup>+</sup>] and [Sb<sup>+</sup>] progeny and establish stocks

♀ y w/y w ; +/TM6, Ubx X ♂ y w/Y; +/+

OR

♂ y w/Y ; +/TM6, Ubx X ♀ y w/y w; +/+

25 B. Schemes used to jump the enhancerless Gal4 insert

1. Jumps from the X-chromosome

♀♀ FM3/FM7, w; +/+ X ♂♂ y w/Y;  $\Delta 2-3$ , Sb/TM6, Ubx

♀♀ FM7, w/ P[Gal4, w<sup>+</sup>] X ♂♂ FM7/Y ;  $\Delta 2-3$ , Sb/+

♀♀ FM7, w/ P[Gal4, w<sup>+</sup>];  $\Delta 2-3$ , Sb/+ X ♂♂ y w/Y; +/+

♀ FM7, w/ y w ;  $\Delta 2-3$ , Sb/+ X ♂ y w/Y; +/+

30 Select [w<sup>+</sup>] and [B] progeny and establish stocks

2. Jumps from the  $\Delta 2-3$ -chromosome

♀♀ y w/y w X ♂♂ y w/Y ; P[Gal4, w<sup>+</sup>],  $\Delta 2-3$ , Sb/+

- 49 -

S lect [w<sup>+</sup>] and [Sb<sup>+</sup>] progeny and establish stocks.

5           C. Chromosomal segregation

To analyze the segregation of the insertions two stocks are used:  
w;+/+; TM3, Sb/CxD and w; CyO/Sco;+/+.

10           Method

To create a large number of strains that express Gal4 in a cell- or tissue-specific manner enhancer detection vectors have been built that carry different versions of the Gal4 gene. Two genes, encoding either the full-length protein or a truncated protein, have been cloned into rosy (ry<sup>+</sup>) and white (w<sup>+</sup>) P-element vectors (modified versions of p1ArB and p1wB; Wilson et al., *Genes and Development* 3:1301-1313 (1989)). Using ry<sup>+</sup> or w<sup>+</sup> as a screen, these vectors have been mobilized by introduction of the A2-3 gene (Robertson et al., *Genetics* 118:461-470 (1988)). To visualize the expression pattern of Gal4, the Gal4 insertion lines are crossed to a strain that carries the lacZ gene under the control of the Gal4 UAS (Fischer et al. *Nature* 332:853-865 1988). Embryos, larvae and adults derived from these crosses are screened for  $\beta$ -galactosidase expression either by an enzyme assay, with X-gal as a substrate, or by staining with monoclonal antibodies against  $\beta$ -galactosidase.  $\beta$ -galactosidase encoded by the UAS-lacZ construct is localized in the cytoplasm.

25           Approximately 500 Gal4-insertion strains have been screened and many that can be used to activate genes in specific tissues have been identified such as, for example, epidermal stripes, mesoderm, the central nervous system and the peripheral nervous system. Many 30 of the lines express  $\beta$ -galactosidase in the salivary

- 50 -

glands as well as in other tissues. It is possible that in constructing the enhancerless-Gal4 transposon a position-dependent salivary gland enhancer was fortuitously generated.

5

## VI. Sample Screen

	# of Strains	No Staining	Salivary Gland	Other Tissues	%
10	9	+	-	-	5.8
	45	-	+	-	28.8
	81	-	+	+	51.9
	21	-	-	+	13.5
	156				100.0

15 To activate a gene (Gene X) in a specific pattern, a Gal4 insertion line is selected and crossed to a strain that carries Gene X cloned behind the GAL UAS.

## 20 VII. Summary of the GAL4/UAS System without the

### Ribozyme

25 The Gal4/UAS system is a two-part system for controlling gene activation. The method is versatile, can be tissue-specific and does not appear to exhibit a basal level of expression except perhaps, as described herein, for a UAS-DTA construct. It can be used to ectopically express characterized genes, to express modified genes that would otherwise be lethal to the organism and to express genes from other species to study their effect on *Drosophila* development. Since 30 the method makes it possible to produce dominant, gain-of-function mutations, epistasis tests and screens for enhancers or suppressors of visible or lethal phenotypes can be carried out. The Gal4 system also 35 allows the expression of toxic products to study the consequences of cell- and tissue-specific ablation.

- 51 -

VIII. Use of Gal4-Expressing Drosophila with the DTA Ribozyme of the Invention

5 Expression of the fusion ribozyme carrying the sequences encoding the DTA protein was placed under the control of a the GAL4 UAS (upstream activator sequence) in pUAST (Figure 10). As stated *supra*, using modified P-element enhancer-trap vectors described above, a  
10 large number of stable lines of *Drosophila* were constructed which each express the yeast transcriptional activator GAL4 in specific spatial and temporal patterns in the developing flies. Any gene under the control of the GAL4 upstream activator sequence (UAS) can be transformed and maintained  
15 singly, then induced in particular *Drosophila* tissues by genetic crossing to lines which express GAL4 (Figure 7). However, it was not possible to take advantage of the Gal4 system for expression of DTA *per se* without further modification, due to the difficulty in  
20 producing UAS-DTA transformants through leaky expression of the DTA.

25 It was found that use of this two-element system as a means of conditionally expressing DTA via a trans-splicing ribozyme (Figure 6) overcame these problems. In those cells expressing GAL4, the GAL4 protein provides the activity necessary for ribozyme transcription, and the GAL4 mRNA provides the target for trans-splicing necessary for DTA production.

30 *Drosophila* embryos may be injected with ribozyme sequences placed under the control of a UAS promoter as described above, using techniques known in the art. Embryos injected with the Rz-DTA<sub>met</sub> construction will not survive, whereas normal transformed flies were obtained from embryos injected with both Rz-DTA<sub>ile</sub> and Rz-DTA<sub>leu</sub>. This result suggested that the internal AUG codon was indeed acting as an initiation codon for the

- 52 -

translation of a toxic product after injection. The codon is adjacent to proposed NAD<sup>+</sup> binding site in the DTA sequence, and to sequences conserved in the distantly related exotoxin A, another EF-2 specific ADP-ribosylase from *Pseudomonas aeruginosa*.

5 ADP-ribosylase from *Psuedomonas aeruginosa*

Transgenic flies containing the Rz-DTA<sub>ile</sub> and Rz-DTA<sub>leu</sub> sequences under control of the GAL4 UAS were crossed to flies producing GAL4 in particular patterns of expression. For example, in one characterized line, 10 line 1J3, the GAL4 gene was been inserted near the hairy gene, and mirrored its pattern of expression. The hairy gene product is produced in epidermal stripes 15 in the even-numbered abdominal segments during embryogenesis. When a UAS-driven LacZ gene was introduced into 1J3 in which GAL4 is expressed in the same pattern as the hairy gene product,  $\beta$ -galactosidase was found localized within the even-numbered stripes. 20 When flies containing the Rz-DTA<sub>leu</sub> gene were crossed to this GAL4-expressing line, normal progeny resulted. However, when flies containing Rz-DTA<sub>ile</sub> were crossed 25 to the GAL4-expressing line, development of the progeny was arrested in embryogenesis. Darker colored bands were evident on the cuticles of the embryos, consistent with the death of underlying cells. When cuticle preparations were examined, the even-numbered denticle bands were disrupted or missing, particularly those of the 4th, 6th and 8th stripes (Figure 11). Other specific patterns of cell death were observed when the 30 containing Rz-DTA<sub>ile</sub> flies are crossed to different GAL4 expressing genes.

- 53 -

Example 5  
Design of Pro-ribozymes

5           As a test for the design of pro-ribozymes, the  
CAT-LacZ trans-splicing ribozyme which described  
earlier was modified (Figure 2). Phylogenetic  
comparisons and mutational analysis (for review, see  
Cech, *Ann Rev. Biochem.* 59:543-568 (1990)) have  
indicated that a core region of the group I self-  
splicing introns is highly conserved and important for  
activity (Figure 8). For the construction of trans-  
splicing pro-ribozymes a helix immediately adjacent to  
this region, P8, was disrupted. In the first  
10          experiments, 13 or 18 nucleotides of new sequence were  
introduced into the 5' strand and loop of helix P8, to  
produce pro-ribozyme 1 and 2, respectively. The extra  
nucleotides were complementary to the 5' "anti-sense"  
15          portion of the ribozyme, while the flanking sequences  
were adjusted to conserve (1) the actual sequences at  
the base of P8, and (2) the extent of base-pairing  
possible within P8 (Figure 13). The extent of self-  
complementarity between the sequences inserted into  
20          helix P8 and the 5' "anti-sense" region of the pro-  
ribozyme is such that this new helix would be expected  
to form in nascent transcripts, in preference to helix  
P8. The formation of this alternative helix would also  
be expected to disrupt flanking secondary and perhaps  
25          tertiary interactions within the catalytic core of the  
ribozyme. Thus, mis-folding of the pro-ribozyme would  
render it catalytically inactive (Figure 14). However,  
base-pairing of the pro-ribozyme with the intended  
target RNA would displace the P8-"anti-sense" base-  
pairing, sequester the "anti-sense" sequences and allow  
30          re-formation of the P8 helix and an active catalytic  
35          region.

- 54 -

domain. Displacement of the P8-''anti-sense'' hairpin results in a greater sum of base-pairs and allows proper folding of the catalytic domain, so should be energetically favored.

5

### CAT-LacZ pro-ribozymes

10

15

20

25

30

35

Cloned sequences corresponding to the two CAT-LacZ pro-ribozymes were constructed using PCR-mutagenesis as discussed above, and RNAs were produced by *in vitro* transcription. The CAT-LacZ trans-splicing ribozyme was observed to undergo scission during transcription at the 3' splice junction, as a result of hydrolysis catalyzed by the intron sequences. Similar hydrolysis is seen in *in vitro* transcripts of the unmodified *Tetrahymena thermophila* intron. In contrast, transcripts of the different CAT-LacZ pro-ribozymes are more stable, with little cleavage evident under the same conditions (Figure 15). This indicates that the pro-ribozymes are inactive, which would be expected if the catalytic sequences were mis-folded. Truncated forms of the pro-ribozymes were tested for specific endoribonuclease activity directed against the CAT RNA. CAT-LacZ pro-ribozyme RNAs were transcribed from templates truncated at the ScaI site, to remove the 3' splice junction and LacZ sequences. Both ribozyme and pro-ribozyme RNAs are stable after removal of the 3' splice site. Incubation of the truncated pro-ribozymes with CAT RNA led to specific cleavage of the target RNA to give fragments of the expected sizes (Figure 16). Specific cleavage activity was seen at 37, 45 and 50 degrees.

Pro-ribozyme forms of the GAL4-DTA trans-splicing ribozyme were also constructed (Figure 17). Regions of 20 nucleotides (complementary to the ''anti-sense'' region) were inserted into the 5' strand and loop of

- 55 -

helix P8. The two pro-ribozymes differed in the extent of base-pairing possible in the modified helices P8, and GAL4-DTA pro-ribozyme 1 possessing both a longer stem and fewer (3) accessible bases in the loop. The helix P8 of GAL4-DTA pro-ribozyme 2 more closely resembles that of the CAT-LacZ pro-ribozyme 2, with a larger loop (14 bases) containing sequences complementary to the "anti-sense" region. Transcripts of the GAL4-DTA pro-ribozymes are more stable than those of the unmodified ribozyme. In particular, pro-ribozyme 2 is mainly intact after incubation in conditions that result in essentially complete self-cleavage of the ribozyme form (30'@ 50°C, 10 mM MgCl<sub>2</sub>, 2 mM GTP, see Figure 18).

15

Having now fully described the invention, it will be understood by those with skill in the art that the scope may be performed within a wide and equivalent range of conditions, parameters and the like, without affecting the spirit or scope of the invention or any embodiment thereof.

20

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: Haseloff, James  
Goodman, Howard M.  
Brand, Andrea  
Perrimon, Norbert

(ii) TITLE OF INVENTION: Cell Ablation Using Trans-Splicing Ribozymes

(iii) NUMBER OF SEQUENCES: 56

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
(B) STREET: 1225 Connecticut Avenue, N.W., Suite 300  
(C) CITY: Washington  
(D) STATE: DC  
(E) COUNTRY: USA  
(F) ZIP: 20036

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: PCT  
(B) FILING DATE: herewith  
(C) CLASSIFICATION:

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 07/642,330  
(B) FILING DATE: 17-JAN-1991

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Goldstein, Jorge A  
(B) REGISTRATION NUMBER: 29,021  
(C) REFERENCE/DOCKET NUMBER: 0609.3496604

## (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (202)833-7533  
(B) TELEFAX: (202)833-8716

## (2) INFORMATION FOR SEQ ID NO :1:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 517 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear
- (ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TGACGCAATT CAACCAAGCG CGGGTAAACG GCGGGAGTAA CTATGACTCT	50
CTAAATAGCA ATATTTACCT TTGGAGGGAA AAGTTATCAG GCATGCACCT	100
CCTAGCTAGT CTTAAACCA ATAGATTGCA TCGGTTAAA AGGCAAGACC	150
GTCAAATTGC GGGAAAGGGG TCAACAGCCG TTCAGTACCA AGTCTCAGGG	200
GAAACTTGA CATGGCCTTG CAAAGGGTAT GGTAATAAGC TGACGGACAT	250
GGTCCTAACCC ACGCAGCCAA GTCTTAAGTC AACAGATCTT CTGTTGATAT	300
GGATGCAGTT CACAGACTAA ATGTCGGTCG GGGAAAGATGT ATTCTTCTCA	350
TAAGATATAG TCGGACCTCT CCTTAATGGG AGGTAGCGGA TGAATGGATG	400
CAACACTGGA GCCGCTGGGA ACTAATTGT ATGCGAAAGT ATATTGATTA	450
GTGTTGGAGT ACTCGTAAGG TAGCCAAATG CCTCGTCATC TAATTAGTGA	500
CGCGCATGAA TGGATTA	517

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 82 bases
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: both
  - (D) TOPOLOGY: linear
- (ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GGCCAAGCTT CTTTACGATG CCATTGGGAT ATATCAACCG TGGTATAAAC	50
CCGTGGTTT TAAAAGTTAT CAGGCATGCA CC	82

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GATTAGTTT GGAGTACTCG TACGGATTCA CGGCCGTCGT TTTACAA

47

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 43 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGCCGAATTTC TTACAATTTC CATTCAAGCT GCGCAACTGT TGG

43

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 623 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGGAGACCGG AAGCTTCTTT ACGATGCCAT TGGGATATAT CAACGGTGGT

50

ATAAAGCCGT GGTTTTAAA AGTTATCAGG CATGCACCTG GTAGCTAGTC

100

TTTAAACCAA TAGATTGCAT CGGTTTAAAA GGCAAGACCG TCAAATTGCG

150

GGAAAGGGGT CAACAGCCGT TCAGTACCAA GTCTCAGGGG AAACTTTGAG

200

ATGGCCTTGC AAAGGGTATG GTAATAAGCT GACGGACATG GTCCTAACCA

250

CGCAGCCAAG TCCTAAAGTCA ACAGATCTTC TGTTGATATG GATGCAGTTC	300
ACAGACTAAA TGTCCGGTCGG GGAAGATGTA TTCTTCTCAT AAGATATAAGT	350
CGGACCTCTC CTTAATGGGA GCTAGC9GAT GAAGTGTGTC AACACTGGAG	400
CCGCTGGAA CTAATTTGTA TGC9AAAGTA TATTGATTAG TTTTGGAGTA	450
CTCGTACGGA TTCACTGGCC GTCGTTTAC AACGTCGTGA CTGGGAAAAC	500
CCTGGCGTTA CCCAACTTAA TCGCCTTGCA GCACATCCCC CTTTCGCCAG	550
CTGGCGTAAT AGCGAAGAGG CCCGCACCGA TCGCCCTTCC CAACAGTTGC	600
GCAGCCTGAA TGGAAATTGT AAG	623

## (2) INFORMATION FOR SEQ ID NO:6:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1038 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

## (11) MOLECULE TYPE: DNA

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GTGACCTTT TTAAGTCGGC AAATATCGCA TGTGTTGTCG ATAGACATCG AGTGGCTTCA	60
AAAGTTATCA GGCATGCACC TGGTAGCTAG TCTTTAAACC AATAGATTGC ATCGGTTAA	120
AAGGCAAGAC CGTCAAATTG CGGGAAAGGG GTCAACAGCC GTTCAGTACC AAGTCTCAGG	180
GGAAACTTG AGATGGCCTT GCAAAGGGTA TGGTAATAAG CTGACGGACA TGGTCCTAAC	240
CACCGAGCCA AGTCCTAAGT CAACAGATCT TCTGTTGATA TGGATGCAGT TCACAGACTA	300
AATGTCGGTC GGGGAAGATG TATTCTTCTC ATAAGATATA GTCGGACCTC TCCTTAATGG	360
GAGCTAGCGG ATGAAGTGAT GCAACACTGG AGCCGCTGGG AACTAATTG TATGCGAAAG	420
TATATTGATT AGTTTTGGAG TACTCGTCTC GATGATGTTG TTGATTCTTC TAAATCTTTT	480
GTGATTGAAA ACTTTCTTC GTACCAACGGG ACTAAACCTG GTTATGTAGA TTCCATTCAA	540

AAAGGTATAAC AAAAGCCAAA ATCTGGTACA CAAGGAAATT ATGACGATGA TTGGAAAGGG	600
TTTTATAGTA CCGACAATAA ATACGACGCT GCGGGATACT CTGTAGATAA TGAAAACCCG	660
CTCTCTGGAA AAGCTGGAGG CGTGGTCAAA GTGACGTATC CAGGACTGAC GAAGGTTCTC	720
GCACACTAAAG TGGATAATGC CGAAACTATT AAGAAAGAGT TAGGTTAAG TCTCACTGAA	780
CCGTTGATGG AGCAAGTCGG AACGGAAGAG TTTATCAAAA CGTTGGTGA TGGTGCTTCG	840
CGTGTAGTGC TCACBCTTCC CTTCGCTGAG GGGAGTTCTA GCCTTGAATA TATTAATAAC	900
TGGGAACAGG CGAAAGCGTT AAGCGTAGAA CTTGAGATTAA ATTTTGAAAC CCGTGGAAAA	960
CGTGGCCAAG ATGCGATGTA TGAGTATATG GCTCAAGCCT GTGCAGGAAA TCGTGTCAAG	1020
CGATCTTTGT GACTCGAG	1038

## (2) INFORMATION FOR SEQ ID NO:7:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 134 bases
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GTTTAGTTGT TCACCTGAGT CGTGTGTTT GTATTTGCG TCTTAGTGTG	50
CCTATGGACA AATCTGGATC TCCCAATGCT AGTAGAACCT CCCGGCGTCG	100
TCGCCCCGCGT AGAGGTTCTC GGTCCGCTTC TGGT	134

## (2) INFORMATION FOR SEQ ID NO:8:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 134 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GTTTAGTTGT TCACCTGAGT CGTGTTCCT TTGTTTGCG TCTCAGTGTG	50
CCTATGGACA AATCTGGATC TCCCAATGCT AGTAGAACCT CCCGGCGTCG	100
TCGCCCGCGT AGAGGTTCTC GGTCCGCTTC TGGT	134

## (2) INFORMATION FOR SEQ ID NO:9:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 152 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GTTATTGTCT ACTGACTATA TAGAGAGTGT TTGTGCTGTG TTTTCTCTTT	50
TGTGTCGTAG AATTGAGTCG AGTCATGGAC AAATCTGAAT CAACCAGTGC	100
TGGTCGTAAC CGTCGACGTC GTCCGCGTCG TGGTTCCCGC TCCGCCCTCCT	150
CC	152

## (2) INFORMATION FOR SEQ ID NO:10:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 152 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GTTATTGTCT ACTGACTATA TAGAGAGTGT GTGTGCTGTG TTTTCTCTTT	50
TGTGTCGTAG AATTGAGTCG AGTCATGGAT AAATCTGAAT CAACCAGTGC	100
TGGTCGTAAC CGTCGACGTC GTCCGCGTCG TGGTTCCCGC TCCGCCCTCCT	150
CC	152

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 131 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AGAGAGTGTG TGTGCTGTGT TTTCTCTTT GTGTCGTAGA ATTGAGTCGA	50
GTCATGGACA AATCTGAATC AACCAGTGCT GGTGTAACC GTCGACGTCG	100
TCCGGGTGCGT GGTTCCCGCT CCGCCCCCTC C	131

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 154 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GTTATTGTCT ACTGATTGTA TAAAGAGTGT GTGTGTGCTG TGTTTTCTCT	50
TTTACGTCGT AGAATTGAGT CGAGTCATGG ACAAACTGA ATCAACCAAGT	100
GCTGGTCGCA ACCGTCGACG TCGTCCGCGT CGTGGTTCCC GCTCCGCCCC	150
CTCC	154

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 154 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(1i) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GTTATTGTCT ACTGACTATA TAGAGAGTGT GTGTGCTG TGTTTCTCT	50
TTTGTGCGT AGAATTGAGT CGAGTCATGG ACAAACTGAA ATCAACCGAGT	100
GCTGGTCGTA ACCGTCGACG TCGTTGCGT CGTGGTCCC GCTCCGCCTC	150
CTCC	154

(2) INFORMATION FOR SEQ ID NO:14:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 130 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GAGTGTGTAT GTGCTGTGTT TTCTCTTTG TGTCGTAGAA TTGAGTCGAG	50
TCATGGACAA ATCTGAATCA ACCAGTGCTG GTCGTAACCG TCGACGTCGT	100
CCCGCGTCGTG GTTCCCGCTC CGCCCCCTCC	130

(2) INFORMATION FOR SEQ ID NO:15:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 152 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GTTATTGTCT ACTGACTATA TAGAGAGTGT GTGTGCTG TGTTTCTCTT	50
TGTGTCGTAG AATTGAGTCG AGTCATGGAC AAATCTGAAT CAACCAAGTGC	100

TGGTCGTAAC CATCGACGTC GTCCGCGTCG TGGTTCCCGC TCCGCCCCCT

150

CC

152

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 78 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GGAGGGGGCG GAGCGGAAAC CACGACGCGG ACGACGTCGA CGGTTACGAC

50

CAGCCCTGGT AGATTCAAGAT TTGTCCAT

78

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 49 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TTTGCCTT AGTGTGCCTA TGGACAAATC TGGATCTCCC AATGCTAGT

49

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 49 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TTTGCCTCTC AGTGTGCCTA TGGACAAATC TGGATCTCCC AATGCTAGT

49

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 56 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TTTGTGTCGT AGAATTGAGT CGAGTCATGG ACAAACTCTGA ATCAACCAGT

50

GCTGGT

56

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 56 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TTTGTGTCGT AGAATTGAGT CGAGTCATGG ATAAAATCTGA ATCAACCAGT

50

GCTGGT

56

## (2) INFORMATION FOR SEQ ID NO:21:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 56 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TTTGTGTCGT AGAATTGAGT CGAGTCATGG ACAAAATCTGA ATCAACCAAGT	50
GCTGGT	56

## (2) INFORMATION FOR SEQ ID NO:22:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 56 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TTTACGTGTCGT AGAATTGAGT CGAGTCATGG ACAAAATCTGA ATCAACCAAGT	50
GCTGGT	56

## (2) INFORMATION FOR SEQ ID NO:23:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 56 bases
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TTTGTGTCGT AGAATTGAGT CGAGTCATGG ACAAAATCTGA ATCAACCAAGT	50
GCTGGT	56

## (2) INFORMATION FOR SEQ ID NO:24:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 56 bases
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TTTGTGTCGT AGAATTGAGT CGAGTCATGG ACAAAATCTGA ATCAACCAGT 50

GCTGGT 56

(2) INFORMATION FOR SEQ ID NO:25:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 56 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TTTGTGTCGT AGAATTGAGT CGAGTCATGG ACAAAATCTGA ATCAACCAGT 50

GCTGGT 56

(2) INFORMATION FOR SEQ ID NO:26:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 59 bases
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AATTTTGTGT CGTAGAATTG AGTCGAGTCA TGGACAAATC TGAATCAACC 50

AGTGCTGCA 59

(2) INFORMATION FOR SEQ ID NO:27:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 51 bases  
 (B) TYPE: Nucleic Acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GCAC TGGTTG ATT CAGATT GTCCATGACT CGACTCAATT CTACGACACA

50

A

51

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 59 bases  
 (B) TYPE: Nucleic Acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

AATTTTGTTG CGTAGAATTG AGTCGAGTCA TGGACAAATC TGAATCAACC

50

AGTGCTGCA

59

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 bases  
 (B) TYPE: Nucleic Acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

AGCATTGGTA TCATCAGGTT TGT

23

## (2) INFORMATION FOR SEQ ID NO:30:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GTTGATGATG TTGTTGATTCT

21

## (2) INFORMATION FOR SEQ ID NO:31:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 amino acids
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

MET ASP LYS PHE ASP ASF ASP VAL VAL ASP SER

5

10

## (2) INFORMATION FOR SEQ ID NO:32:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 bases
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ATGGACAAAT TTGATGATGT TGTTGATTCT

30

## (2) INFORMATION FOR SEQ ID NO:33:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 59 bases  
(B) TYPE: Nucleic Acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AATTTGTGT CGTAAATTG AGTCGAGTCA TGGACAAATC TGAATCAACC

50

AGTGCTGCA

59

## (2) INFORMATION FOR SEQ ID NO:34:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17 bases  
(B) TYPE: Nucleic Acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

AGCCATCCTT GGTCAG

17

## (2) INFORMATION FOR SEQ ID NO:35:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15 bases  
(B) TYPE: Nucleic Acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GTAAGGGTGG ATGTT

15

- 71 -

## (2) INFORMATION FOR SEQ ID NO:36:

## (1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 amino acids  
(B) TYPE: Amino Acid  
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

MET ASP LYS SER GLU LEU ARG VAL ASP VAL

1 5 10

## (2) INFORMATION FOR SEQ ID NO:37:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 bases  
(B) TYPE: Nucleic Acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

ATGGACAAAT CTGAATTAAG GGTGGATGTT

30

## (2) INFORMATION FOR SEQ ID NO:38:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 70 bases  
(B) TYPE: Nucleic Acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TCTCGATGAT GTTGTTGATT CTTCTAAATC TTTTGTGATG GAAAACTTT

50

CTTCGTACCA CGGGACTAAA

70

- 72 -

## (2) INFORMATION FOR SEQ ID NO:39:

## (1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11 amino acids  
(B) TYPE: Amino Acid  
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

NET GLU ASN PHE SER SER TYR HIS GLY THR LYS

1 5 10

## (2) INFORMATION FOR SEQ ID NO:40:

## (1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 70 bases  
(B) TYPE: Nucleic Acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

TCTCGATGAT GTTGTTGATT CTTCTAAATC TTTTGTGATT GAAAACTTTT

50

CTTCGTACCA CGGGACTAAA

70

## (2) INFORMATION FOR SEQ ID NO:41:

## (1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 70 bases  
(B) TYPE: Nucleic Acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

TCTCGATGAT GTTGTTGATT CTTCTAAATC TTTTGTGTTG GAAAACTTTT

50

CTTCGTACCA CGGGACTAAA

70

## (2) INFORMATION FOR SEQ ID NO:42:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 78 bases
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

ATGAAGCTTC TCGATGATGT TGTTGATTCT TCTAAATCTT TTGTGATGGA

50

AAACTTTCT TCGTACCAACG GGACTAAA

78

## (2) INFORMATION FOR SEQ ID NO:43:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 amino acids
- (B) TYPE: Amino Acid
- (D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

MET LYS LEU LEU ASP ASP VAL VAL ASP SER SER SER LYS SER PHE VAL							
1	5	10	15				

MET GLU ASN PHE SER SER TYR HIS GLY THR LYS							
20	25						

## (2) INFORMATION FOR SEQ ID NO:44:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 bases
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

ATGGAGAAAA AAATCACTGG ATATACCACC GTTGATATAT C

41

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17 amino acids  
(B) TYPE: Amino Acid  
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

MET GLU LYS LYS ILE THR ASP SER LEU ALA VAL VAL LEU GLN ARG  
1 5 10 15

ARG ASP

17

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 51 bases  
(B) TYPE: Nucleic Acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

ATGGAGAAAA AAATTACGGA TTCACTGGCC GTCGTTTAC AACGTCGTGA

50

C

51

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 40 bases  
(B) TYPE: Nucleic Acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

ATGAAGCTAC TGTCTTCTAT CGAACAGCA TGGATATTT

40

(2) INFORMATION FOR SEQ ID NO:48:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 amino acids  
(B) TYPE: Amino Acid  
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

MET LYS LEU LEU ASP ASP VAL VAL ASP SER SER LYS SER PHE VAL  
1 5 10 15

MET GLU ASN PHE SER  
20

(2) INFORMATION FOR SEQ ID NO:49:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 bases  
(B) TYPE: Nucleic Acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

ATGAAGCTTC TCGATGATGT TGTTGATTCT TCTAAATCTT TTGTGATGGA

50

AAACTTTCT

60

(2) INFORMATION FOR SEQ ID NO:50:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 72 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

## (11) MOLECULE TYPE: DNA

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:50:

AUGGAGAAAA AAAUCACUGG AUAUACCACC GUUGAUAUAU CCCAAUGGCA UCGUAAAGAA	60
CAUUUUGAGG CA	72

## (2) INFORMATION FOR SEQ ID NO:51:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 479 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:51:

AAGCUUCUUU ACGAUGCCAU UGGGAUAUAU CAACGGUGGU AUAAAGCCGU GGUUUUUAAA	60
AGUUAUCAGG CAUGCACCUG GUAGCUAGUC UUUAAAACCAA UAGAUUGCAU CGGUUUUUAAA	120
GGCAAGACCG UCAAAUUGCG GGAAAGGGGU CAACAGCCGU UCAGUACCAA GUCUCAGGGG	180
AAACUUUGAG AUGGCCUUGC AAAGGGUAUG GUAAUAAGCU GACGGACAUG GUCCUAACCA	240
CGCAGCCAAG UCCUAAGUCA ACAGAUCUUC UGUUGAUUAUG GAUGCAGUAC AGACUAAAUG	300
UCGGUCGGGG AAGAUUGUAUU CUUCUCAUAA CAUAUAGUCG GACCUUCUCCU UAAUGGGAGC	360
UAGCGGAUGA AGUGAUGCAA CACUGGAGCC GCUGGGAACU AAUUUGUAUG CGAAAGUAUA	420
UUGAUUAGUU UUGGAGUACU CGUACGGAUU CACUGGCCGU CCUGUUACAA CGUCGUGAC	479

## (2) INFORMATION FOR SEQ ID NO:52:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 479 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: both

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

AAGCUUCUUU ACGAUGCCAU UGGGAUUAU CAACGGUGGU AUAAAGCCGU GGUUUUUAAA	60
AGUUAUCAGG CAUGCACCUG GUAGCUAGUC UUUAAACCAA UAGAUUGCAU CGGUUUAAAA	120
GGCAAGACCG UCAAAUJUGCG GGAAAGGGGU CAACAGCCGU UCAGUACCAA GUCUCAGGGG	180
AAACUUUGAG AUGGCCUUGC AAAGGGUAUG GUAAUAAGCU GACGGACAUG GUCCUAACCA	240
CGCAGCCAAG UCCUAAGUCA ACAGAUCUUC UGUUGAUUAUG GAUGCAGUAC AGACUAAAUG	300
UCGGUCGGGG AAGAUGUAUU CUUCUCAUAA CAUAUAGUCG GACCUCUCCU UAAUGGGAGC	360
UAGCGGAUGA AGUGAUGCAA CACUGGAGCC GCUGGGAACU AAUUGUAUG CGAAAGUAUA	420
UUGAUUAGUU UUGGAGUACU CGUACGGAUU CACUGGCCGU CCUGUUACAA CGUCGUGAC	479

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 480 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

AAGCUUCUUU ACGAUGCCAU UGGGAUUAU CAACGGUGGU AUAAAGCCGU GGUUUUUAAA	60
AGUUAUCAGG CAUGCACCUG GUAGCUAGUC UUUAAACCAA UAGAUUGCAU CGGUUUAAAA	120
GGCAAGACCG UCAAAUJUGCG GGAAAGGGGU CAACAGCCGU UCAGUACCAA GUCUCAGGGG	180
AAACUUUGAG AUGGCCUUGC AAAGGGUAUG GUAAUAAGCU GACGGACAUG GUCCUAACCA	240
CGCAGCCAAG UCCUAAGUCA ACAGAUCUUC UGUUGAUUAUG GAUGCAGUAC AGACUAAAUG	300

UCGGUCGGGA CCGUUGAUAU AUGGUUCAUA ACAUAUAGUC GGACCUCCUCC UUAAUGGGAG	360
CUAGCGGAUG AAGUGAUGCA ACACUGGAGC CGCUGGGAAC UAAUJUGUAU GCGAAAGUAU	420
AUJUGAUUAGU UUUGGGAGUAC UCGUACGGAU UCACUGGCCG UCCUGUUACA ACGUCGUGAC	480

## (2) INFORMATION FOR SEQ ID NO:54:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 487 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

AAGCUUCUUU ACGAUGCCAU UGGGAUUAU CAACGGUGGU AUAAAGCCGU GGUUUUUAAA	60
AGUUAUCAGG CAUGCACCUUG GUAGCUAGUC UUUAAACCAA UAGAUJUGCAU CGGUUUAAAA	120
GGCAAGACCG UCAAAUJUGCG GGAAAGGGGU CAACAGCCGU UCAGUACCAA GUCUCAGGGG	180
AAACUUUGAG AUGCUCUUGC AAAGGGUAUG GUAAAAGGU GACGGACAUG GUCCUAACCA	240
CGCAGCCAAG UCCUAAGUCA ACAGAUCUUC UGUUGAUUAUG GAUGCAGUAC AGACUAAAUG	300
UCGGUCGGGA CCGUUGAUAU AUCCCAAACG GUUCAUAACA UAUAGUCGGA CCUCUCUUA	360
AUGGGAGCUA GCGGAUGAAG UGAUGCAACA CUGGAGCCGC UGGGAACUAA UUUGUAUGCG	420
AAAGUAUAUU GAUJAGUJUUU GGAGUACUCG UACGGAUUCA CUGGCCGUCC UGUUACAACG	480
UCGUGAC	487

## (2) INFORMATION FOR SEQ ID NO:55:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1044 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

GTCGACCTT TTAAGTCGGC AAATATCGCA TGTTTGTTCG ATAGACATCG AGTGGCTTCA	60
AAAGTTATCA GGCATGCACC TGGTAGCTAG TCTTAAACC AATAGATTGC ATCGGTTAA	120
AAGGCAAGAC CGTCAAATTG CGGGAAAGGG GTCAACAGCC GTTCAGTACC AAGTCTCAGG	180
GGAAACTTG AGATGGCCTT GCAAAGGGTA TGGTAATAAG CTGACGGACA TGGCCTAAC	240
CACGCAGCCA AGTCCTAAGT CAACAGATCT TCTGTTGATA TGGATGCAGT TCACAGACTA	300
AATGTCGGTC GGGGAACAAC ATGCGATATT GTTCTCATAA GATATAGTCG GACCTCTCCT	360
TAATGGGAGC TAGCGGATGA AGTGATGCAA CACTGGAGCC GCTGGAACT AATTTGTATG	420
CGAAAGTATA TTGATTAGTT TTGGAGTACT CGTCTCGATG ATGTTGTGA TTCTTCTAAA	480
TCTTTGTGA TTGAAAACCTT TTCTTCGTAC CACGGGACTA AACCTGGTTA TGTAGATTCC	540
ATTCAAAAAG GTATACAAAAA GCCAAAATCT GGTACACAAG GAAATTATGA CGATGATTGG	600
AAAGGGTTTT ATAGTACCGA CAATAAATAC GACGCTGCGG GATACTCTGT AGATAATGAA	660
AACCCGCTCT CTGGAAAAGC TGGAGGC GTCAAAGTGA CGTATCCAGG ACTGACGAAG	720
GTTCTCGCAC TAAAAGTGG A TAATGCCAA ACTATTAAGA AAGAGTTAGG TTTAAGTCTC	780
ACTGAACCGT TGATGGAGCA AGTCGGAACG GAAGAGTTA TCAAAAGGTT CGGTGATGGT	840
GCTTCGCGTG TAGTGCTCAG CCTTCCCTTC GCTGAGGGGA GTTCTAGCGT TGAATATATT	900
AATAACTGGG AACAGGGGAA AGCGTTAACG GTAGAACTTG AGATTAATT TGAAACCCGT	960
GGAAAACGTG GCCAAGATGC GATGTATGAG TATATGGCTC AAGCCTGTGC AGGAAATCGT	1020
GTCAGGCGAT CTTTGTGACT CGAG	1044

## (2) INFORMATION FOR SEQ ID NO:56:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1047 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:56:

GTCGACCTTT TTAAGTCGGC AAATATCGCA TGTTTGTTCG ATAGACATCG AGTGGCTTCA	60
AAAGTTATCA GGCGATGCACC TGGTAGCTAG TCTTTAAACC AATAGATTGC ATCGGTTAA	120
AAGGCAAGAC CGTCAAATTG CGGGAAAGGG GTCAACAGCC GTTCAGTACC AAGTCTCAGG	180
GGAAACTTTG AGATGGCCTT GCAAAGGGTA TGGTAATAAG CTGACGGACA TGGTCCTAAC	240
CACGCAGCCA AGTCCTAAGT CAACAGATCT TCTGTTGATA TGGATGCAGT TCACAGACTA	300
AATGTCGGTC GGGCAAACAT GCGATATTG CCGTTGTCA TAAGATATAG TCGGACCTCT	360
CCTTAATGGG AGCTAGCGGA TGAAGTGATG CAACACTGGA GCCGCTGGGA ACTAATTGT	420
ATGCGAAAGT ATATTGATTA GTTTTGGAGT ACTCGCTCTCG ATGATGTTGT TGATTCTCT	480
AAATCTTTG TGATTGAAAA CTTTCTTCG TACCACTGGGAA CTAACCTGG TTATGTAGAT	540
TCCATTCAA AAGGTATACA AAAGCCAAA TCTGGTACAC AAGGAAATTA TGACGATGAT	600
TGGAAAGGGT TTTATAGTAC CGACAATAAA TACGACGCTG CGGGATACTC TGTAGATAAT	660
GAAAACCCGC TCTCTGGAAA AGCTGGAGGC GTGGTCAAAG TGACGTATCC AGGACTGACG	720
AAGGTTCTCG CACTAAAAGT GGATAATGCC GAAACTATTAA AGAAAGAGTT AGGTTAAAGT	780
CTCACTGAAC CGTTGATGGA GCAAGTCGGA ACGGAAGAGT TTATCAAAAG GTTCGSTGAT	840
GGTGCTTCGC GTGTAGTGCT CAGCCTTCCC TTGCGCTGAGG GGAGTTCTAG CGTTGAATAT	900
ATTAATAACT GGGAACAGGC GAAAGCGTTA AGCGTAGAAC TTGAGATTAA TTTGAAACC	960
CGTGGAAAAC GTGGCCAAGA TGCGATGTAT GAGTATATGG CTCAAGCCTG TGCAGGAAAT	1020
CGTGTCAAGGC GATCTTTGTG ACTCGAG	1047

- 81 -

What is claimed is:

1. A polynucleotide molecule, said molecule encoding a trans-splicing ribozyme, the sequence of said ribozyme being a fusion RNA, the sequence of such fusion RNA comprising:

- (1) a first RNA sequence, said first RNA sequence being sufficient for targeting said ribozyme to hybridize to a target RNA, and
- (2) a second RNA sequence, said second RNA sequence being capable of being transferred colinearly into a target RNA as a result of the trans-splicing activity of said ribozyme;

wherein expression of said polynucleotide molecule is operably linked to expression of a transcriptional activator protein, and wherein said first RNA sequence is a sequence that hybridizes to an RNA that encodes said transcriptional activator protein.

2. The polynucleotide molecule of claim 1, wherein said transcriptional activator is GAL4.

3. The polynucleotide molecule of claim 1, wherein said second RNA sequence comprises a sequence that encodes a peptide toxic to the host cell.

4. The polynucleotide molecule of claim 3, wherein said peptide is the DTA peptide.

5. The polynucleotide molecule of claim 4, wherein said DTA peptide is a mutant peptide sequence.

6. The polynucleotide molecule of claim 5, wherein said mutant peptide sequence comprises amino acids encoded by SEQ ID. No. 40.

- 82 -

7. The polynucleotide molecule of claim 5, wherein said mutant peptide sequence comprises amino acids encoded by SEQ ID. No. 41.

5 8. The polynucleotide molecule of claim 1, wherein said first RNA sequence is a sequence that hybridizes to GAL4 RNA and wherein said second RNA sequence is a sequence that encodes the DTA peptide.

10 9. The polynucleotide molecule of any one of claims 1-8, wherein said molecule is RNA.

15 10. The polynucleotide molecule of any one of claims 1-8, wherein said molecule is DNA.

15 11. A polynucleotide molecule comprising a ribozyme expression cassette, said cassette being capable of being stably inserted into the genome of a host, and said cassette comprising the sequence of a promoter capable of functioning in such host, operably-linked to the coding sequence of the polynucleotide of any one of claims 1-9.

25 12. A host cell comprising the polynucleotide molecule of claim 11.

13. The host cell of claim 12, wherein said host cell is a viral cell.

30 14. The host cell of claim 12, wherein said host cell is a prokaryotic cell.

15. The host cell of claim 12, wherein said host cell is a eukaryotic cell.

- 83 -

16. The host cell of claim 15, wherein said eukaryotic cell is a plant cell.

5 17. The host cell of claim 15, wherein said eukaryotic cell is an animal cell.

10 18. The host cell of claim 17, wherein said animal is *Drosophila*.

15 19. The host cell of claim 17, wherein said animal is a mammal.

20 20. The host cell of claim 19, wherein said animal is human.

15 21. A method for *in vitro* trans-splicing, such method comprising the steps of:

20 22. A method for *in vivo* trans-splicing, said method comprising the steps of:

(1) providing the polynucleotide molecule of claim 9 in a trans-splicing reaction mixture, said polynucleotide comprising a sequence capable of hybridizing with a second polynucleotide;

(2) providing said second polynucleotide to such reaction mixture; and

25 (4) catalyzing the trans-splicing of said second polynucleotide under said conditions.

22. A method for *in vivo* trans-splicing, said method comprising the steps of:

30 (1) providing a polynucleotide of claim 9 to a host cell;

(2) expressing said ribozyme encoded by said molecule in said host cell;

(3) expressing a substrate of said ribozyme in said host cell; and

35 (4) catalyzing the trans-splicing of said ribozyme with such substrate in said host cell.

23. A method for inactivating the activity of a target RNA, said method comprising:

5 (1) providing the polynucleotide of claim 9 to a trans-splicing reaction mixture, said ribozyme possessing catalytic activity against a target RNA, said catalytic activity resulting in the inactivation of the functioning of said target RNA;

10 (2) providing said target RNA to said mixture; and

(3) providing conditions that allow said polynucleotide to express said catalytic activity.

24. A method for providing a desired genetic sequence to a host cell *in vivo*, said method comprising:

15 (1) providing the polynucleotide of claim 9 to said host cell, said polynucleotide possessing catalytic activity against a target RNA in said host cell, said ribozyme being capable of trans-splicing said desired genetic sequence;

20 (2) providing said target RNA in said host cell;

and

25 (3) providing conditions that allow said ribozyme to trans-splice said desired genetic sequence into the sequence of said target RNA.

26. A method for cell ablation in multicellular plants and animals, said method comprising providing the polynucleotide of claim 9 to a fertilized embryonic host cell, said ribozyme encoding the sequence of a peptide toxic to said host cell and said ribozyme being capable of trans-splicing said sequence into a target in said host cell.

35 27. A method for engineering male or female sterility in a plant, said method comprising providing the polynucleotide of claim 9 to a germ cell of said

sp cies, said ribozyme being targeted to an RNA that, when expressed as a protein, is necessary for the fertility of said plant and results in ablation of the cell that expresses said protein.

5

27. A method for immunizing plants against a plant pathogen, said method comprising transforming plant cells with the polynucleotide of claim 9, wherein said polynucleotide encodes a trans-splicing sequence capable of providing immunity against said pathogen to said plant, and wherein infection of a cell from said plant with said pathogen results in the ablation of said cell.

10

15 28. The polynucleotide molecule of any one of claims 1-7 wherein said ribozyme is a pro-ribozyme.

29. The polynucleotide molecule of any one of claims 8 wherein said ribozyme is a pro-ribozyme.

20

30. The polynucleotide molecule of any one of claims 9 wherein said ribozyme is a pro-ribozyme.

25

31. The polynucleotide molecule of any one of claims 10 wherein said ribozyme is a pro-ribozyme.

32. The polynucleotide molecule of claim 11, wherein said ribozyme is a pro-ribozyme.

30

33. The host cell claim 12, wherein said ribozyme is a pro-ribozyme.

34. The method of claim 21, wherein said ribozyme is a pro-ribozyme.

35

35. The method of claim 22, wherein said ribozyme is a pro-ribozyme.

5

36. The method of claim 23, wherein said ribozyme is a pro-ribozyme.

10

37. The method of claim 24, wherein said ribozyme is a pro-ribozyme.

15

38. The method of claim 25, wherein said ribozyme is a pro-ribozyme.

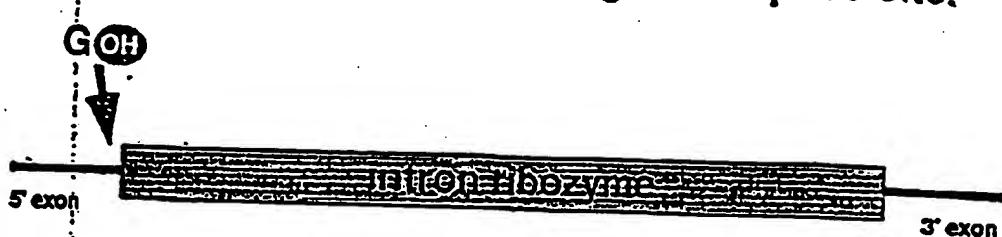
39. The method of claim 26, wherein said ribozyme is a pro-ribozyme.

40. The method of claim 27, wherein said ribozyme is a pro-ribozyme.

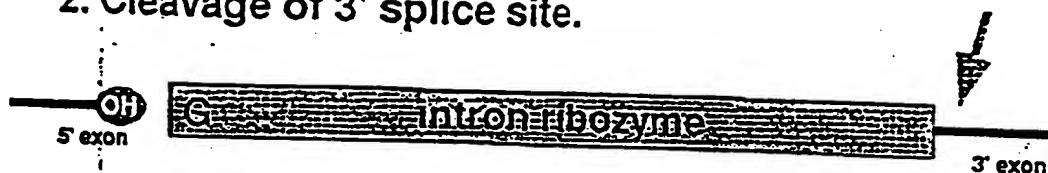
FIGURE 1

## Self-splicing of Group I introns.

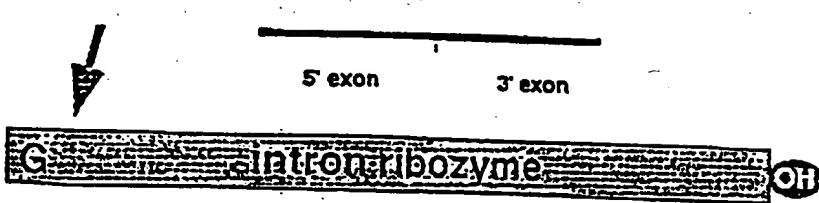
### 1. Guanosine-mediated cleavage of 5' splice site.



### 2. Cleavage of 3' splice site.



### 3. Ligation of exon segments.



### 4. Circularization of intron.

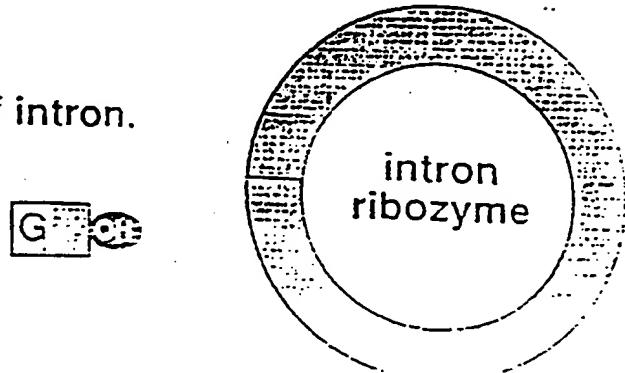
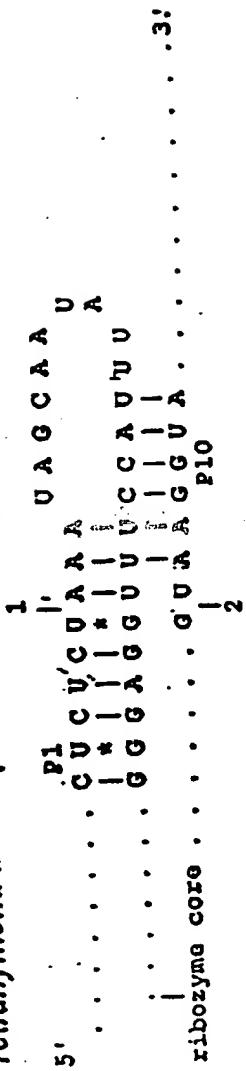


FIGURE 2

Design of ribozymes for trans-splicing.

A)

*Tetrahymena thermophila* self-splicing rRNA Intron



B) Target mRNA and trans-splicing ribozyme

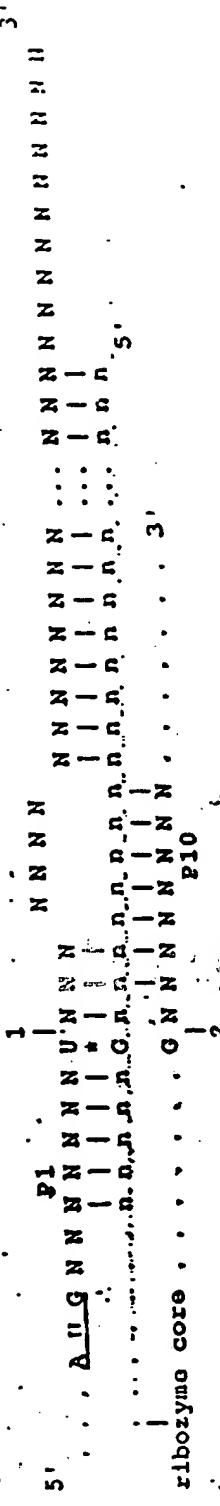
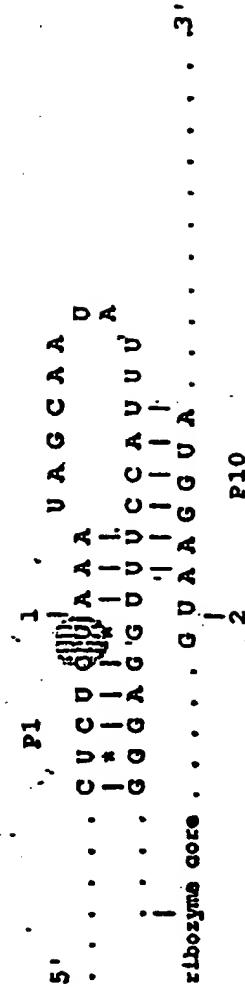


FIGURE 3

## Design of CAT-LacZ $\alpha$ -peptide trans-splicing ribozyme.

## *Tetrahymena thermophila* self-splicing rRNA Intron



## CAT-LacZ trans-splicing ribozyme.

300 15 No.

5'. AUG GAG AAA AAA AGG GAU UCA CGU UUA CGA CGU CGU GAC  
met glu lys lys ile thr asp ser leu ala val val leu gln arg arg asp.

45 46

41 DNA Strider 1.0 623 Today, January 15, 1991 5:07:14 PM

T-1C aCAT-LacZ ts-R2 -> 11st

NA sequence : 623 b.p. : CGGAGACCGGA ... CGAAATTGTAAC linear

1 10	1 20	1 30	1 40	1 50	1 60
1 GGGAGACCGG	AACTTCTT	ACGATCCAT	TCGGATATAT	CACCGCTCT	ATAAACCCCT
61 GGTTTTAA	AGTTATCAGG	CATGCCACCTG	GTAACGATCT	TTTAACCAA	TAGATGCCAT
121 CGGTTTAA	CGGACAGCT	TCAATTGCG	GGAAAGGGT	CACAGCCCT	TCACACCAA
181 GTCAGAGG	AACTTCTGAG	AAGGCCCTGC	AAACCTTAAC	GGAAATAGCT	CACCCACATG
241 GTCCTAACCA	CGAACCCAC	TCCCTAAGTC	ACAGATCTTC	TGTTTGTATG	GTGCGAGTC
301 ACAGACTAA	TGTCGGCTG	GGAGAGCTGA	TTCTTCAT	AGATATAACT	CGGCGCTTC
361 CTTAATGGG	GCACCGCT	GAACGTGATCC	AAACCTGGAG	CGCGTGGGA	CTAAATGTA
421 TCCGAACTA	TTTGTGAA	TTTGGCTCTA	CTGGTACGAA	TTCTCTGCC	CTCTTTTAC
481 AACGCTGAA	CTGGGAAAC	CTTCGGCTTA	GGCAACTTA	GGCGCTTGA	GCACATCCCC
541 CTTCGGGAC	CTGGGCTAA	ACCGAACGAG	GGCGACCGA	TGCCCCCTTC	CAACACTTC
601 CGAACCGAA	CGGAAATGCT	AACGAGACGG			

1 10 1 20 1 30 1 40 1 50 1 60 623

(SEQ ID No. 5)

FIGURE 3(b)

## Cucumber mosaic virus RNA 4 trans-splicing rib zymes.

## Virus RNA target sequences.

SEQ ID NO

### Oligonucleotide target sequences.

## CMV RNA 4 - diphtheria toxin A-chain trans-splicing ribozymes.

### ‘bozyme 1

15024mu2

AAUUTUDGUGUCGUAGAUAUGAGUCAGUCAGGACAAATCUGAAU.CAACCGAGUCUGCA 3'	33
.....     *   .....    (46 bp)	
.....GACUUG GUUCCUACCGA ... 5'	34
.....*     *	
RIBOZYME CORE.....G.UAAGGGUGGGAUGUU ... 3'	35

met asp lys ser gln leu arg val asp val 36  
 AUG GAC AAA UCU GAA UUA AGG GUG GAU GUU ... 3' 37

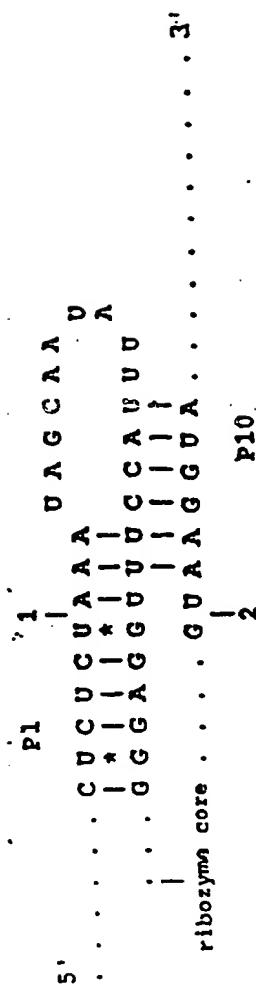
## COMPARISON OF CUCUMBER MOSAIC VIRUS RNA 3' S SEQUENCES.

THE CPO FTS ARE SET FOR PCT OCA CTA CPO 115 001 SET OCA ECA OCA ... CTA AND OCA ARE CPO AND PCT  
CONFIDENTIAL INFORMATION

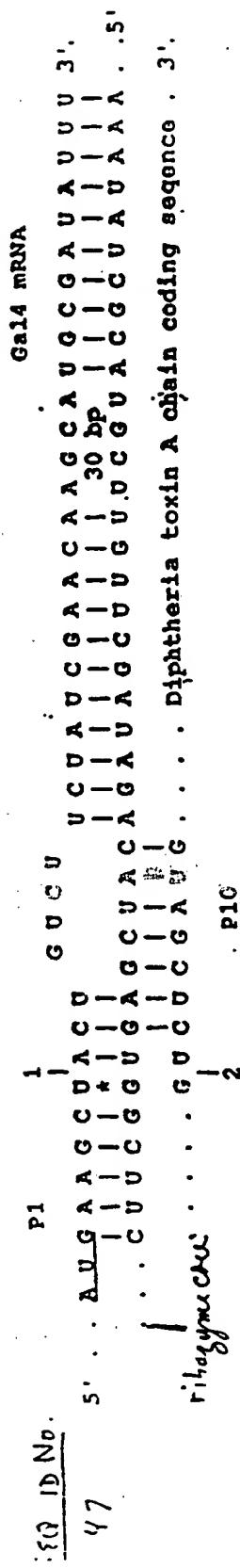
## Figure 5

## FIGURE 6

## Design of Gal4 - Diphtheria toxin A-chain trans-splicing ribozyme. *Tetrahymena thermophila* self-splicing rRNA intron.



## Galt·Dir·A trans-splicing ribozyme.



\*\*\* DNA Strider \*\*\* Tuesday, January 15, 1991 5:09:45 PM

TT-aCal6-DT tsRz MUTANT - List

DNA sequence 1038 b.p. CTGGACCTTTT TTCTGACTCCAG linear

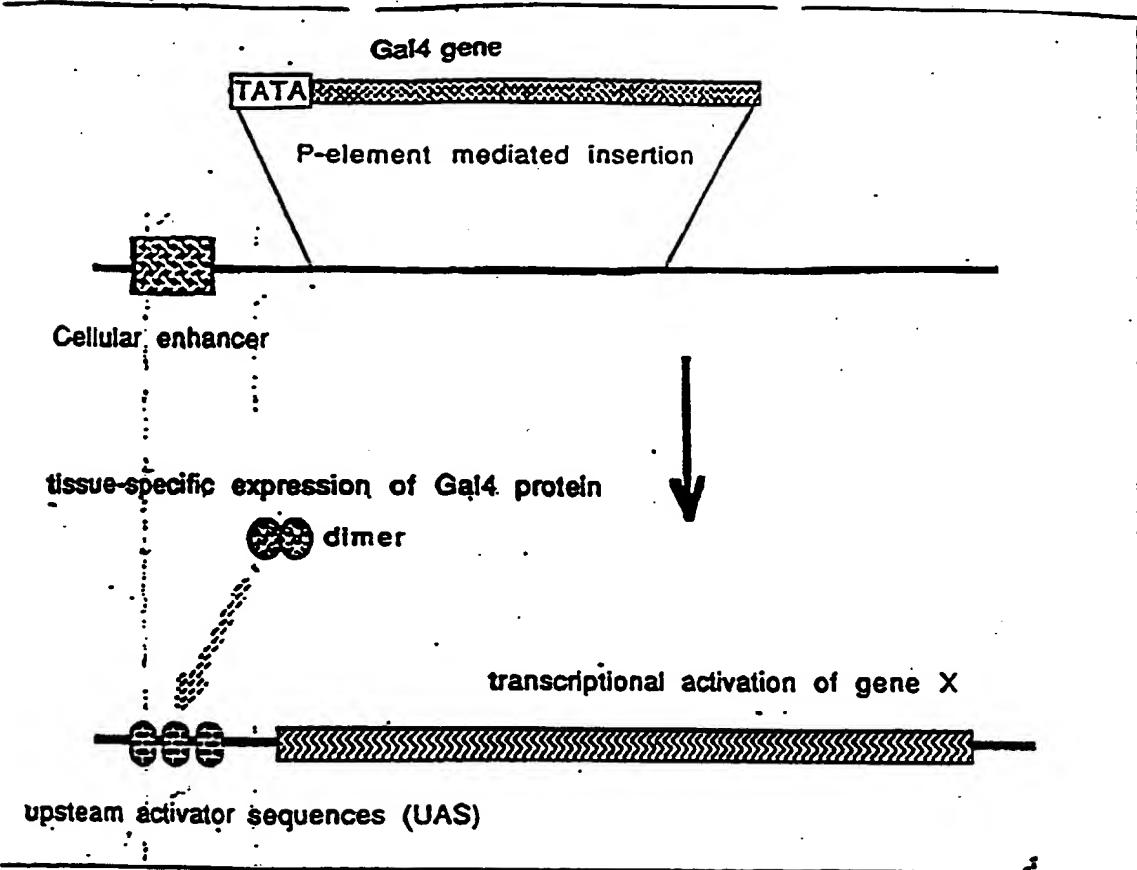
1	10	19	20	30	40	50	60
1	CTGGACCTTT	TTAAGTCGGC	AAATATCGCA	TCCTTGTGTC	AAAGACATCC	ACTGGCTTC	CA
61	AAAGCTATCA	GGCATTCGCC	TGGTACCCAG	TCTTAAACC	AAAGACATCC	ACTGGCTTC	60
121	AAGCCGAGAC	CTGTCAGAATC	CGGGAAAGCC	CTGAAACGCG	TGGTAAAGAC	CTGACGAGC	120
181	CGAAACCTTC	AGATGGCTT	GCNAAGCTA	TCTCTGAAAG	TGGTAAAGAC	CTGACGAGC	180
241	CACCCACCCA	AGCTCTAAGT	CAACAGCTC	TCTCTGAAAG	TGGTAAAGAC	CTGACGAGC	240
301	AAATGTCGTC	GGGGAAAGCA	TATTCCTC	TCTCTGAAAG	TGGTAAAGAC	CTGACGAGC	300
361	GAGCTTACGGC	ATGAACTGAT	GCCTACCTGG	ATACAGATA	GTGGACGCTC	TCCCTAACTG	360
421	TATATTCG	AGCTTGGAC	TACTCTCTC	ACCCGCTGG	ACCTGAACTTC	TATGGGAAAG	420
481	CGCATGAA	AGCTTCTTC	CTGGCCCGC	GTGAGTCCTC	TCATTCCTTC	TCATTCCTTC	480
541	AAAGGTTAAC	AAAGGCGAAA	ATCTGCTACA	ACCTTACCG	TCATTCCTTC	TCATTCCTTC	540
601	TTTAACTTA	CGGCAATAA	AAACGAGCTC	CGGGAAATT	ATGGGATCA	TGGAAAGGG	600
661	CTCTCTGGAA	AACTGGAGG	CTGGCTTAA	CGGGAAACTC	CTGAGAAATA	TCGAAACCGC	660
721	GCACGAAA	TCCTATTCG	CGAAACCTAT	GTGAGCTATC	CAGGCTTCAC	GAAGCTTC	720
781	CCCTTGGC	AGCTTCTGG	AAACGAGAG	AAAGAAGAGT	TCCTTTCAGC	TCATTCCTTC	780
841	CTCTGATTC	TCACCCCTCC	CTTCCCGAG	TTTACAAA	GTGGCTGCA	TCATTCCTTC	840
901	TGGGAAACGG	CGAAACCTA	AAACGAGAGA	GGCTGTTCA	TCCTTCAATA	TATTAATAC	900
961	CTGGCCGAG	AGCTGATCTA	TGAGTAAATG	CTTGAGATA	ATTTCGAAAC	CGCTGAAATA	960
1021	CGATCTTGT	GAATCGAC		GTCTCACCTT	GTGACCAA	TCTCTCAGG	1020

1038

SEQ ID No. ?

FIGURE 6(3)

FIGURE 7



P-element mediated "enhancer-trapping" for expression of Gal4 protein.

10 / 20

## Prevention of toxin expression from splicing by-products.

## translation of trans-spliced mRNA

translation of 3' "exons"

5' -  
TCGCTATCTCTGATCTCTTAACCTTCAG  
Arg Gln Asp Ser Tyr His Gly Thr Lys  
Met Glu Asn Phe Ser Ser Tyr His Gly Thr Lys  
58  
39

## Figure 8

FIGURE 9

GAL4 VECTOR (pGaTB or pGaTN)

(Unique sites in italics; 3' NotI site is unique in pGaTB only)

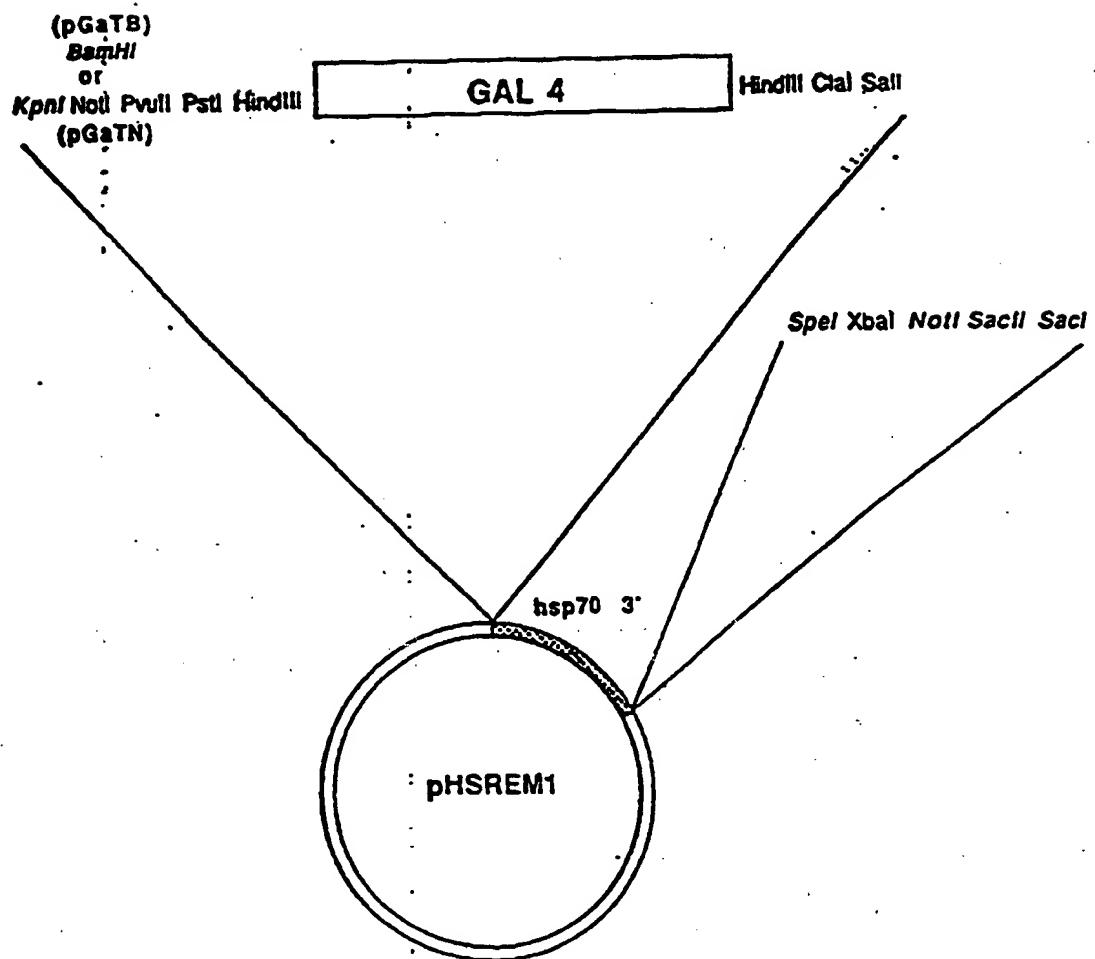


FIGURE 10  
GAL UAS VECTOR (pUAST)  
(Unique sites in italics)

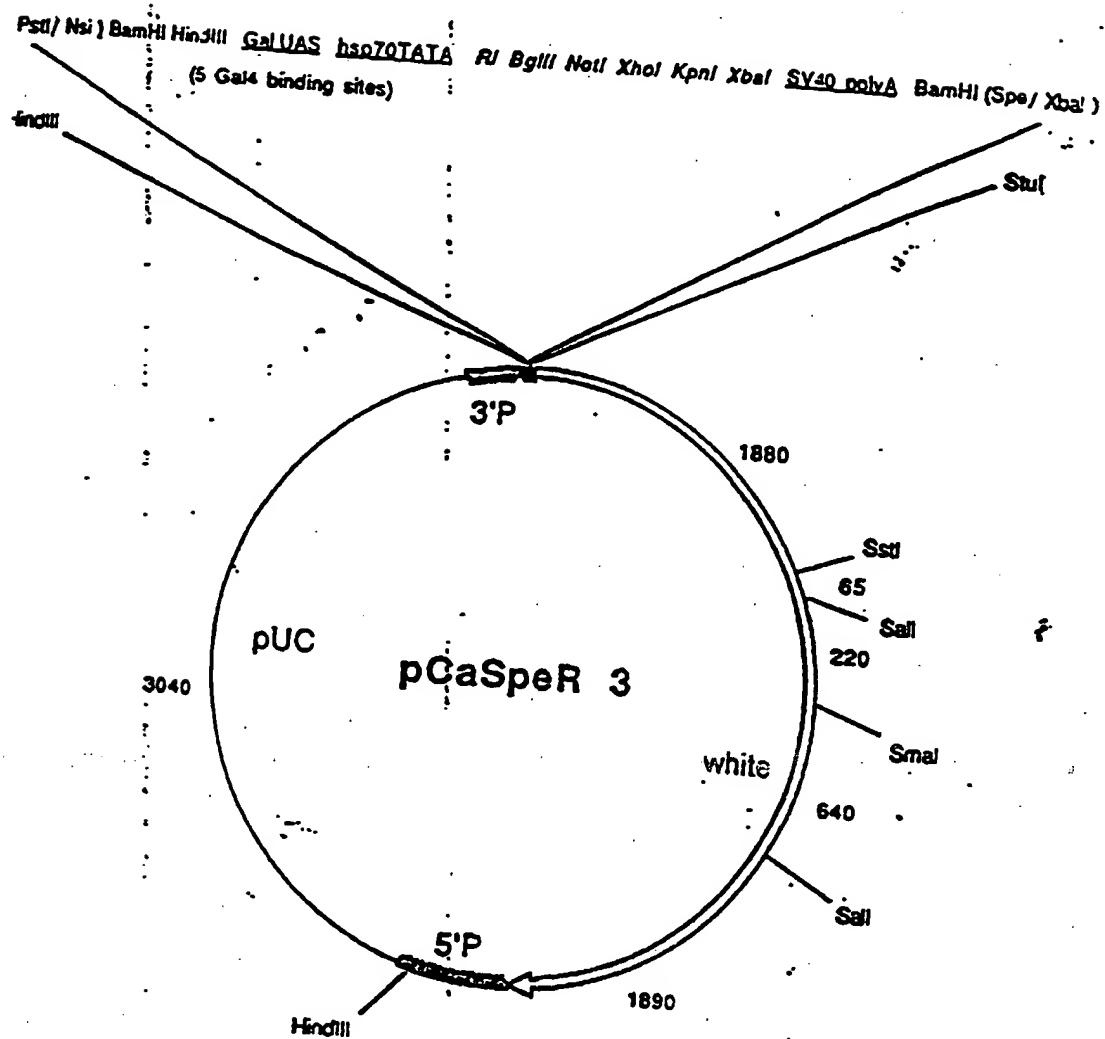
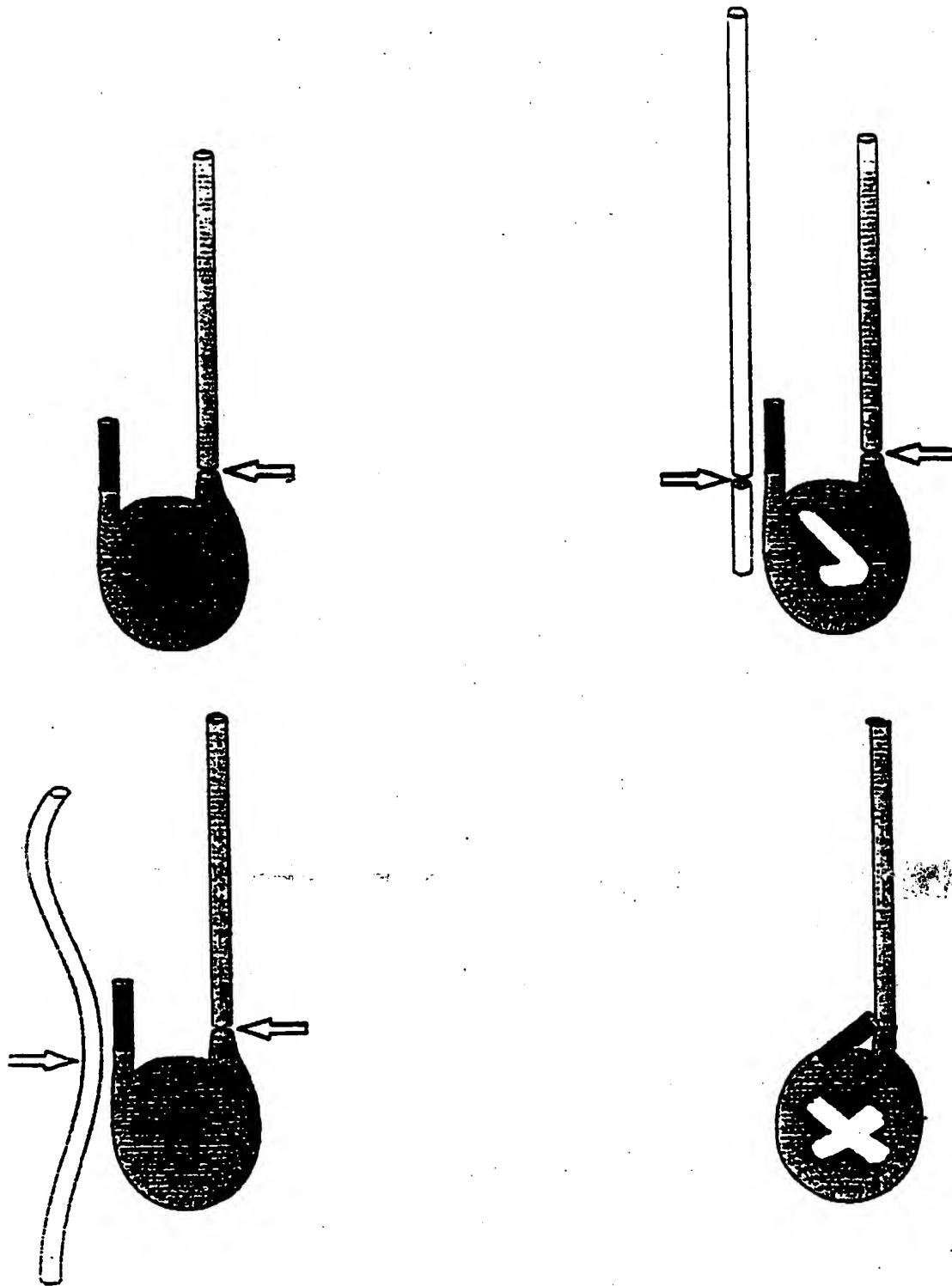




FIGURE 11

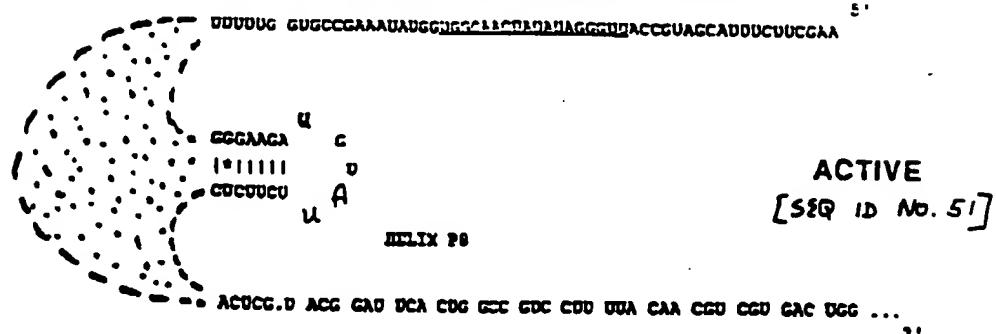
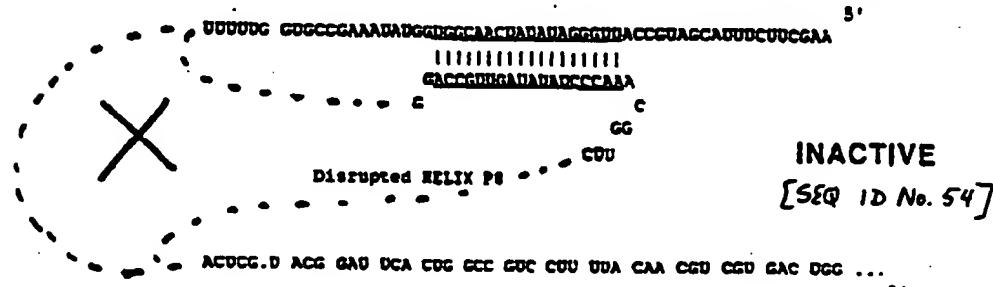
FIGURES 11





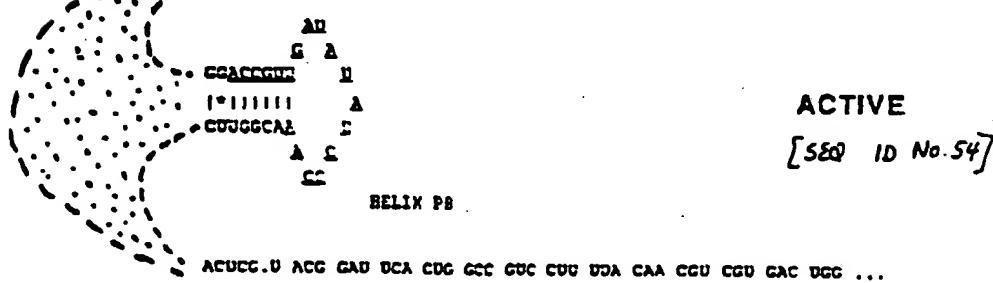
## FIGURE 14

## "Pro-ribozyme" design.

(1). CAT-LacZ  $\alpha$ -peptide trans-splicing ribozyme(2). CAT-LacZ  $\alpha$ -peptide trans-splicing "pro-ribozyme" #2

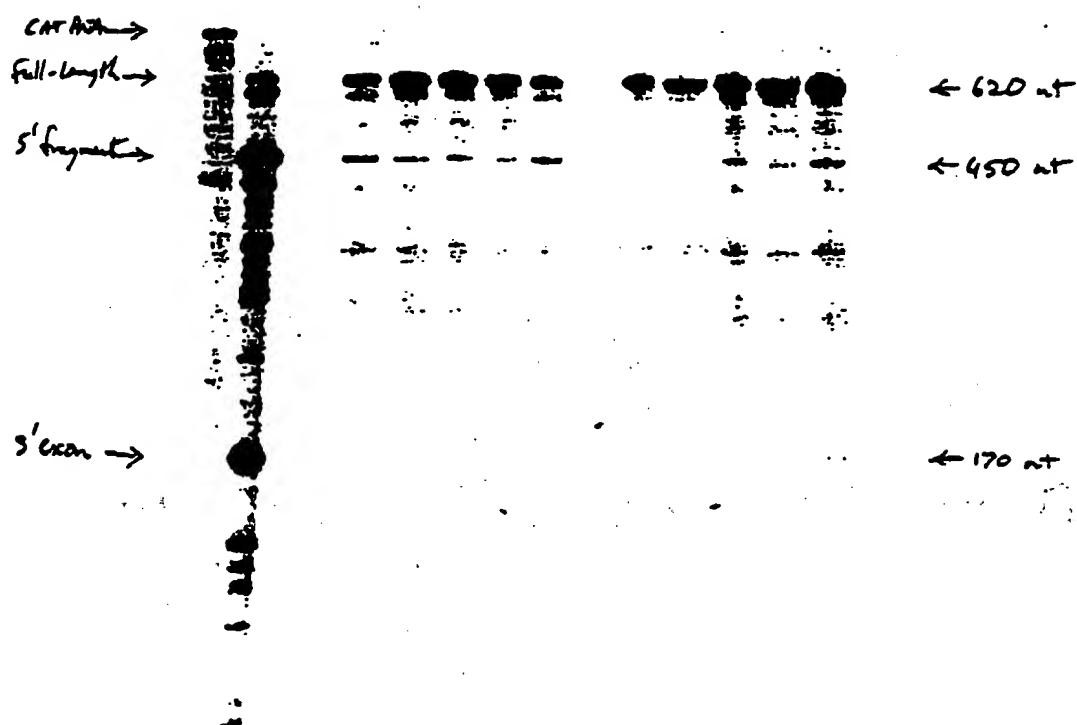
5' ..AUGGAGAAAAAAAC CACUG GAAUACCCACCGGGAGAGADCCCACGGCAUCGGAAAGAACAUUUCGACCA..... 3'

5' ..UUUUGG GUGCCGAAAUAGGGGGCAACGAAUAGGGGGACCGUACCCAUCCUUCUCCAA 3'



## FIGURE 15

CAT - Lac Z

R<sub>2</sub> pro-R<sub>2</sub>-1 pro-R<sub>2</sub>-2

CAT-Lac Z

FIGURE 16

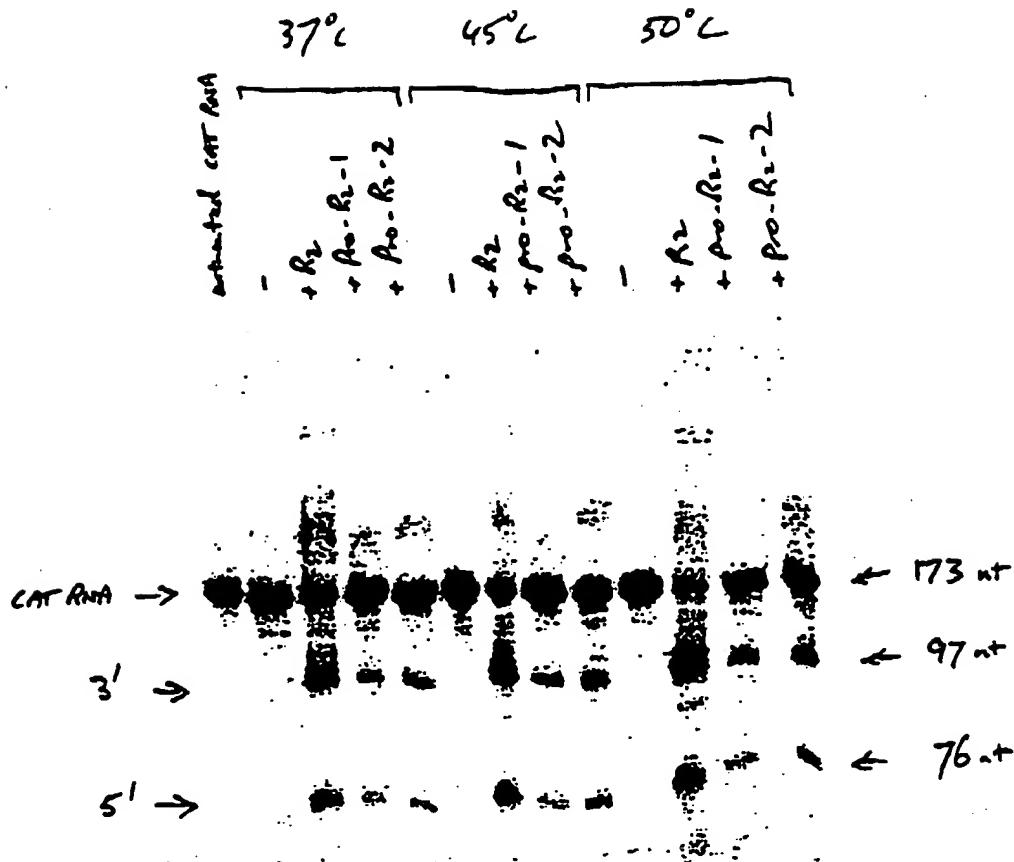


FIGURE 17

## Modification of helix P8 for pro-ribozyme construction.

## Wild-type

[SEQ ID No. 51]

```

G G G A A G A   U
| * | | | * |   G
C U C U U C U   U
                  A
                  U

```

## GAL4-DTA pro-ribozyme 1

[SEQ ID No. 55]

```

G G G A A C A A   C   G   20 nt
| * | | | | | |   A U
C U C U U G U U   U A
                  A   G

```

## CAT-LacZ pro-ribozyme 1

[SEQ ID No. 53]

```

G G A C C G U   U G   13 nt
| * | | | * |   A
C U U G G U A   U
                  U A

```

## CAT-LacZ pro-ribozyme 2

[SEQ ID No. 54]

```

A U A
G G A C C G U U   G   U   16 nt
| * | | | | | |   A
C U U G G C A A   A
                  C C C   U

```

## GAL4-DTA pro-ribozyme 2

[SEQ ID No. 56]

```

U G C G
A   A   20 nt
G G C A A A C   C
| * | | | | |   A
C U G U U U G   U
                  C G U U

```

N = complementary to ribozyme "anti-sense"

INCREASING STABILITY



- FIGURE 18 -

GA24-DT-A

Diagram of a polymer chain segment GA24-DT-A showing three repeating units: pro-R<sub>2</sub>-2, pro-R<sub>2</sub>-1, and R<sub>2</sub>.

$(\text{pro-})R_2 \rightarrow$   $\leftarrow 1050 \text{ nT}$

$\begin{matrix} 3' \\ 5' \end{matrix} \rightarrow$   $\leftarrow 600 \text{ nT}$

$\leftarrow 450 \text{ nT}$

50°C, 60°

# INTERNATIONAL SEARCH REPORT

International Application No. PCT/US92/00277

<b>I. CLASSIFICATION OF SUBJECT MATTER</b> (if several classification symbols apply, indicate all) <sup>3</sup>		
According to International Patent Classification (IPC) or to both National Classification and IPC		
IPC (S) : Please See Attached Sheet. US CL : 435/91, 172.1, 172.3, 193, 240.2, 240.4, 252.3; 536/27; 514/44		
<b>II. FIELDS SEARCHED</b>		
Minimum Documentation Searched <sup>4</sup>		
Classification System	Classification Symbols	
U.S.	435/91, 172.1, 172.3, 193, 240.2, 240.4, 252.3; 536/27; 514/44	
Documentation Searched other than Minimum Documentation to the extent that such Documents are included in the Fields Searched <sup>5</sup>		
MEDLINE, EMBASE, WPI, DIALOG search terms: ribozyme, trans, splice, intron, DTA, selective, lethal, insect, drosophila, antisense		
<b>III. DOCUMENTS CONSIDERED TO BE RELEVANT<sup>14</sup></b>		
Category <sup>6</sup>	Citation of Document <sup>15</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No. <sup>18</sup>
X/Y	Nature, Volume 322, issued 03 July 1986, J.W. Szostak, "Enzymatic Activity of the Conserved Core of a Group I Self-Splicing Intron," pages 83-86, see Figure 4.	1,9,10,21, 23,28,30,31,34 ,36/1-40
X/Y	Nature, Volume 322, issued 03 July 1986, Garriga et al., "Mechanism of Recognition of the 5' Splice Site in Self-splicing Group I Introns," pages 86-89, see Figure 1.	1,9,10,21, 23,28,30,31,34 ,36/1-40
X/Y	Gene, Volume 82, issued 1989, G.F. Joyce et al., "Amplification, Mutation, and Selection of Catalytic RNA," pages 83-87, see Figure 3.	1,9,10,21, 23,28,30,31,34 ,36/1-40
X/Y	Nature, Volume 342, issued 23 November 1989, Michel et al., "The Guanosine Binding Site of the Tetrahymena Ribozyme," pages 391-395, see Figure 4(a).	1,9,10,21, 23,28,30,31,34 ,36/1-40
X/Y	Proceedings of the National Academy of Science, Volume 86, issued October 1989, Coudra et al., "RNA Structure, Not Sequence, Determines the 5' Splice-Site Specificity of a Group I Intron," pages 7402-7406, see Figure 1.	1,9,10,21, 23,28,30,31,34 ,36/1-40
<p>* Special categories of cited documents:<sup>16</sup></p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&amp;" document member of the same patent family</p>		
<b>IV. CERTIFICATION</b>		
Date of the Actual Completion of the International Search <sup>2</sup>	Date of Mailing of this International Search Report <sup>2</sup>	
10 APRIL 1992	08 MAY 1992	
International Searching Authority <sup>1</sup>	Signature of Authorized Officer <sup>20</sup>	
ISA/US	Richard Lebovitz	

## FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

X/Y	EMBO Journal, Volume 9, No. 6, issued June 1990, Winter et al., "The Mechanism of Group I Self-Splicing: An Internal Guide Sequence Can Be Provided In Trans," pages 1923-1928, see Figure 5.	1,9,10,21, 23,28,30,31,34 ,36/1-40
X/Y	Molecular and Cellular Biology, Volume 10, No. 6, issued June 1990, Suh et al., "Base Pairing Between The 3' Exon and an Internal Guide Sequence Increases 3' Splice Site Specificity in the Tetrahymena Self-Splicing rRNA Intron," pages 2960-2965, see Figure 1.	1,9,10,21, 23,28,30,31,34 ,36/1-40

V.  OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE<sup>1</sup>

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1.  Claim numbers , because they relate to subject matter (1) not required to be searched by this Authority, namely:

2.  Claim numbers , because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out (1), specifically:

3.  Claim numbers , because they are dependent claims not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI.  OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING<sup>2</sup>

This International Searching Authority found multiple inventions in this international application as follows:

Please See Attached Sheet.

- As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application. (Telephone Practice)
- As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:
- No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:
- As all searchable claims could be searched without effort justifying an additional fee, the International Search Authority did not invite payment of any additional fee.

Remark on protest

The additional search fees were accompanied by applicant's protest.  
 No protest accompanied the payment of additional search fees.

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)		
Category <sup>1</sup>	Citation of Document, <sup>18</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No. <sup>18</sup>
X/Y	Journal of Molecular Biology, Volume 211, issued 1990, Salvo et al., "Deletion-Tolerance and Trans-splicing of the Bacteriophage T4 td Intron," pages 537-549, see Figure 6.	1,9,10,21, 23,28,30,31,34 ,36/1-40
Y	Cell, Volume 40, issued February 1985, Waring et al., "The Tetrahymena rRNA Intron Self-Splices in E. Coli: In Vivo Evidence for the Importance of Key Base-Paired Regions of RNA for RNA Enzyme Function," pages 371-380, see Table 1.	12,14,33
Y	Cancer Research, Volume 46, issued 1986, I.H. Maxwell et al., "Regulated Expression of a Diphtheria Toxin A-chain Gene Transfected Into Human Cells: Possible Strategy for Inducing Cancer Cell Suicide," pages 4660-4664, see abstract EMBASE No. 86247841.	3-8
Y	Proceedings of the National Academy of Science, Volume 87, issued 1990, T. Yagi et al., "Homologous Recombination at c-fyn Locus of Mouse Embryonic Stem Cells With Use of Diphtheria Toxin A-fragment Gene in Negative Selection," pages 9918-9922, see abstract EMBASE No. 91036311.	3-8
Y	Proceedings of the National Academy of Science, Volume 87, No. 13, issued 1990, C.J. Riedel et al., "Diphtheria Toxin Mutant Selectively Kills Cerebellar Purkinje Neurons," pages 5051-5055, see abstract EMBASE No. 90230135.	3-8
Y	WO, A, 90/13654 (Grill et al.) 15 November 1990, see Abstract.	11,13,15,16,22 ,24-27,32, 35,37-40
Y	EP, A, 0,240,208 (Shewmaker et al.) 07 October 1987, see Abstract.	11,13,15,16,22 ,24-27,32, 35,37-40
Y	WO, A, 89/05852 (Haseloff et al.) 29 June 1989, see pages 9-11.	11,13,15,16,22 ,24-27,32, 35,37-40
Y	WO, A, 87/03451 (Neiman et al.) 18 June 1987, see page 3.	17,19,20
Y	Proceedings of the National Academy of Science, Volume 85, issued 1988, Qian et al., "Antisense Ribosomal Protein Gene Expression Specifically Disrupts Oogenesis in Drosophila melanogaster," pages 9601-9605, see Abstract.	18

FURTHER INFORMATION CONTINUED FROM PREVIOUS SHEETS

I. CLASSIFICATION OF SUBJECT MATTER:

IPC (5):

C12P 19/34; C12N 5/10, 7/01, 9/10, 9/12, 15/00, 15/10, 15/82, 15/85; C07H 15/12

VI. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

I. Claims 1-21 and 28-35, drawn to polynucleotides, in vitro methods of trans-splicing, and host cells, classified in 435/172.3, 240.2, 240.4.

II. Claim 22, drawn to an in vivo method of trans-splicing, classified in 435/91.

III. Claims 23 and 36, drawn to an in vivo method of inactivating target RNA, classified in 435/172.1, 172.3

IV. Claims 25, 26, 38, and 39, drawn to a method of cell ablation in plants and animals and methods for engineering male or female sterility, classified in 435/172.3.

V. Claims 24 and 37, drawn to another method of in vivo trans-splicing, classified in class 435/91, 172.3.

VI. Claims 27 and 40, drawn to a method of immunizing plants against a plant pathogen, classified in 435/235.1, 172.3.

Inventions I and any of II, III, IV and VI are related as product and additional process of use. The inventions are distinct because the polyoligonucleotide product can be used in a materially difference process of using that product such as for a hybridization probe.

Each of Inventions II through VI are independent and distinct inventions because they can be practiced separately and independently of each other and each recite different process steps.

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET		
X/Y	EMBO Journal, Volume 9, No. 6, issued June 1990, Winter et al., "The Mechanism of Group I Self-Splicing: An Internal Guide Sequence Can Be Provided In Trans," pages 1923-1928, see Figure 5.	1, 9, 10, 21, 23, 28, 30, 31, 34 , 36/1-40
X/Y	Molecular and Cellular Biology, Volume 10, No. 6, issued June 1990, Suh et al., "Base Pairing Between The 3' Exon and an Internal Guide Sequence Increases 3' Splice Site Specificity in the Tetrahymena Self-Splicing rRNA Intron," pages 2960-2965, see Figure 1.	1, 9, 10, 21, 23, 28, 30, 31, 34 , 36/1-40
<b>V. <input type="checkbox"/> OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE<sup>1</sup></b> <p>This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:</p> <ol style="list-style-type: none"> <li><input type="checkbox"/> Claim numbers .., because they relate to subject matter (1) not required to be searched by this Authority, namely:</li> <li><input type="checkbox"/> Claim numbers .., because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out (1), specifically:</li> <li><input type="checkbox"/> Claim numbers .., because they are dependent claims not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).</li> </ol>		
<b>VI. <input checked="" type="checkbox"/> OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING<sup>2</sup></b> <p>This International Searching Authority found multiple inventions in this international application as follows: Please See Attached Sheet.</p> <ol style="list-style-type: none"> <li><input checked="" type="checkbox"/> As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application. (Telephone Practice)</li> <li><input type="checkbox"/> As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:</li> <li><input type="checkbox"/> No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:</li> <li><input type="checkbox"/> As all searchable claims could be searched without effort justifying an additional fee, the International Search Authority did not invite payment of any additional fee.</li> </ol> <p>Remark on protest</p> <p><input type="checkbox"/> The additional search fees were accompanied by applicant's protest. <input checked="" type="checkbox"/> No protest accompanied the payment of additional search fees.</p>		

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)		
Category <sup>*</sup>	Citation of Document, <sup>16</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No. <sup>18</sup>
X/Y	Journal of Molecular Biology, Volume 211, issued 1990, Salvo et al., "Deletion-Tolerance and Trans-splicing of the Bacteriophage T4 td Intron," pages 537-549, see Figure 6.	1, 9, 10, 21, 23, 28, 30, 31, 34, 36/1-40
Y	Cell, Volume 40, issued February 1985, Waring et al., "The Tetrahymena rRNA Intron Self-Splices in E. Coli: In Vivo Evidence for the Importance of Key Base-Paired Regions of RNA for RNA Enzyme Function," pages 371-380, see Table 1.	12, 14, 33
Y	Cancer Research, Volume 46, issued 1986, I.H. Maxwell et al., "Regulated Expression of a Diphtheria Toxin A-chain Gene Transfected Into Human Cells: Possible Strategy for Inducing Cancer Cell Suicide," pages 4660-4664, see abstract EMBASE No. 86247841.	3-8
Y	Proceedings of the National Academy of Science, Volume 87, issued 1990, T. Yagi et al., "Homologous Recombination at c-fyn Locus of Mouse Embryonic Stem Cells With Use of Diphtheria Toxin A-fragment Gene in Negative Selection," pages 9918-9922, see abstract EMBASE No. 91036311.	3-8
Y	Proceedings of the National Academy of Science, Volume 87, No. 13, issued 1990, C.J. Riedel et al., "Diphtheria Toxin Mutant Selectively Kills Cerebellar Purkinje Neurons," pages 5051-5055, see abstract EMBASE No. 90230135.	3-8
Y	WO, A, 90/13654 (Grill et al.) 15 November 1990, see Abstract.	11, 13, 15, 16, 22, 24 - 27, 32, 35, 37-40
Y	EP, A, 0,240,208 (Shewmaker et al.) 07 October 1987, see Abstract.	11, 13, 15, 16, 22, 24 - 27, 32, 35, 37-40
Y	WO, A, 89/05852 (Haseloff et al.) 29 June 1989, see pages 9-11.	11, 13, 15, 16, 22, 24 - 27, 32, 35, 37-40
Y	WO, A, 87/03451 (Neiman et al.) 18 June 1987, see page 3.	17, 19, 20
Y	Proceedings of the National Academy of Science, Volume 85, issued 1988, Qian et al., "Antisense Ribosomal Protein Gene Expression Specifically Disrupts Oogenesis in Drosophila melanogaster," pages 9601-9605, see Abstract.	18

FURTHER INFORMATION CONTINUED FROM PREVIOUS SHEETS

I. CLASSIFICATION OF SUBJECT MATTER:  
IPC (5):

C12P 19/34; C12N 5/10, 7/01, 9/10, 9/12, 15/00, 15/10, 15/82, 15/85; C07H 15/12

VI. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING  
This ISA found multiple inventions as follows:

I. Claims 1-21 and 28-35, drawn to polynucleotides, in vitro methods of trans-splicing, and host cells, classified in 435/172.3, 240.2, 240.4.

II. Claim 22, drawn to an in vivo method of trans-splicing, classified in 435/91.

III. Claims 23 and 36, drawn to an in vivo method of inactivating target RNA, classified in 435/172.1, 172.3

IV. Claims 25, 26, 38, and 39, drawn to a method of cell ablation in plants and animals and methods for engineering male or female sterility, classified in 435/172.3.

V. Claims 24 and 37, drawn to another method of in vivo trans-splicing, classified in class 435/91, 172.3.

VI. Claims 27 and 40, drawn to a method of immunizing plants against a plant pathogen, classified in 435/235.1, 172.3.

Inventions I and any of II, III, IV, V and VI are related as product and additional process of use. The inventions are distinct because the polyoligonucleotide product can be used in a materially difference process of using that product such as for a hybridization probe.

Each of Inventions II through VI are independent and distinct inventions because they can be practiced separately and independently of each other and each recite different process steps.